The Dissertation Committee for Lindsey Janae Smith certifies that this is the approved version of the following dissertation:

A Comparison of Procedures for Handling Missing School Identifiers with the MMREM and HLM

Committee:

S. Natasha Beretvas, Supervisor

Barbara Dodd

Matt Hersh

Keenan Pituch

Tiffany Whittaker

A Comparison of Procedures for Handling Missing School Identifiers with the MMREM and HLM

by

Lindsey Janae Smith, B.S.Bioch.; M.S.Stat.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2012

A Comparison of Procedures for Handling Missing School Identifiers with the MMREM and HLM

Lindsey Janae Smith, Ph.D.
The University of Texas at Austin, 2012

Supervisor: S. Natasha Beretvas

This simulation study was designed to assess the impact of three ad hoc procedures for handling missing level two (here, school) identifiers in multilevel modeling. A multiple membership data structure was generated and both conventional hierarchical linear modeling (HLM) and multiple membership random effects modeling (MMREM) were employed. HLM models purely hierarchical data structures while MMREM appropriately models multiple membership data structures. Two of the ad hoc procedures investigated involved removing different subsamples of students from the analysis (HLM-*Delete* and MMREM-*Delete*) while the other procedure retained all subjects and involved creating a pseudo-identifier for the missing level two identifier (MMREM-*Unique*). Relative parameter and standard error (*SE*) bias were calculated for each parameter estimated to assess parameter recovery. Across the conditions and parameters investigated, each procedure had some level of substantial bias. MMREM-*Unique* and MMREM-*Delete* resulted in the least amount of relative parameter bias while HLM-*Delete* resulted in the least amount of relative *SE* bias. Results and implications for applied researchers are discussed.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Multilevel modeling is commonly used to handle the dependency resulting from clustered data (e.g., Raudenbush & Bryk, 2002). Clustered data result from a dataset that consists, for example, of outcomes for individuals who share a common context that result in some degree of similarity or dependence in the individuals. Two examples of two-level clustered data are as follows: students (level one unit) nested in schools (level two unit) and patients (level one unit) nested in therapy groups (level two unit). Use of the conventional hierarchical linear model (HLM) involves the assumption that each level one unit is a member of only one level two unit, but this is not always the case. When this pure hierarchy is not present, rather some or all of the level one units are members of more than one level two unit, a multiple membership data structure is present (Goldstein, 2010).

There are many reasons why a multiple membership data structure will arise. In the medical field, patients could be members of one or more therapy groups. In an educational context, students can be members of multiple schools. When a student is a member of multiple elementary schools (for example), a multiple membership data structure is present. In this case, multiple schools contribute to the academic achievement of a student and should be accounted for. Conventional HLM assumes each level one unit is a member of a single level two unit and thus cannot be used to handle a multiple membership data structure. Multiple membership random effects modeling (MMREM; Goldstein, 2010; Rasbash & Browne, 2001) can, however, handle this type of data structure.

Previous methodological research has investigated MMREM. Some of the studies have compared parameter estimates using both MMREM and conventional HLM when a multiple membership data structure is present. Researchers found that the MMREM estimates tend to be superior to the conventional HLM estimates (Chung & Beretvas, 2012; Wolff Smith & Beretvas, 2011). Another study investigated the impact of weight assignment with MMREM and concluded that the choice of weights did not greatly impact parameter estimation (Wolff Smith & Beretvas, 2012). While not simulation studies, Goldstein, Burgess, and McConnell (2007) and Leckie (2009) have compared MMREM with HLM estimates using real data. Both studies found the variance of the highest level of data was underestimated when HLM was used to model the data as compared with MMREM.

All of the simulation study research has been conducted assuming an ideal scenario in which there are no missing level two identifiers. Additionally, mobility was randomly assigned. The scope of an applied study is typically limited to a set of a few schools. Thus, if a student moves into or out of that particular sample of schools, then the student would have at least one missing school identifier. Hill and Goldstein (1998) presented a way to handle such missing identifiers for scenarios in which level one units are nested within two higher level classifications where some identifiers of one classification are missing. The Hill and Goldstein procedure assigns a set of weights for possible units of the unknown classification. While the authors demonstrated their procedure using a two-level cross-classified dataset, the two higher level classifications do not have to be cross-classified. However, each level one unit must have at least two higher level classifications (e.g., elementary and middle school) associated with it. The current study will focus on a two-level scenario in which there is only one higher level classification and in which the Hill and Goldstein procedure could not be applied.

In addition to not addressing the issue of missing level two units, previous MMREM methodological research has not assessed differences between MMREM and conventional HLM when mobility status is authentically assigned. Results have indicated that MMREM and HLM fixed effects are very similar even when the dataset being analyzed with HLM consisted only of non-mobile students' data. When simulating mobility, however, previous studies have randomly assigned mobility status. It is far more likely that mobility is an endogenous predictor of student outcomes such as student achievement. Thus, mobility is not randomly assigned, rather it is a function of other unmodeled student characteristics.

This study has been designed to address these two deficiencies in previous MMREM methodological research. The current study will investigate three ad hoc procedures for handling missing school identifiers in multiple membership data, namely HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*. The HLM-*Delete* procedure matches what is frequently done in applied research. Using this procedure, all students who are mobile will be removed from the analysis. The MMREM-*Delete* procedure will remove students' data if they are missing at least one school identifier. Under the MMREM-*Unique* procedure, a unique, place-holding school identifier will be substituted for each missing school identifier. Additionally, mobility will be generated as a function of student characteristics that will be unmodeled in the estimating MMREM so as to mimic real data more accurately than has been done in the past. Four design conditions will be manipulated in this study, namely, the percent of mobile students, the intra-class correlation coefficient, the number of level two units (here, schools), and the percent of mobile students missing a level two identifier (here, school). The fixed effects and random effects variance component estimates will be summarized across conditions and ad hoc procedures (HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*). The relative parameter and standard error bias (Hoogland & Boomsma, 1998) will then be calculated to assess which ad hoc procedure performs best.

Chapter 2: Literature Review

The foci of this study are to investigate methods of handling a missing level two identifier as well as to more realistically generate mobility status. In this chapter, a general introduction to multilevel modeling will be provided. Following that, the conventional hierarchical linear model (HLM) and multiple membership random effects model (MMREM), an extension of the HLM, will be discussed.

MULTILEVEL MODELING

Multilevel modeling is used to appropriately model clustered data. Clustered datasets are those in which individuals (level one units) share common contexts (level two units). This introduces a dependency. Examples of such datasets can be found in many fields. For example, in education, students are nested in schools; in business, associates are nested in firms; and in medicine, patients are nested in therapy groups. Individuals within each cluster (e.g., school) will be more similar to one another than an individual in another cluster. Conventional single-level models are not designed to handle this type of dependency. A single-level regression model can be presented as follows:

$$Y_i = \beta_0 + \beta_1 X_i + e_i \tag{1}$$

where Y_i is the outcome for individual i, β_0 is the mean outcome for a person with X_i equal to zero, β_1 is the slope of the regression line, X_i is individual i's value on the independent variable, and e_i is the error (or residual) for individual i which is assumed normally and *independently* distributed with a mean of zero and a constant variance of σ^2. When handling clustered data, the independence assumption is violated potentially resulting in an increase in the type I error rate due to the underestimation of the standard errors (Fielding & Goldstein, 2006). Thus, this model (Equation 1) cannot be used when clustered data are present. Alternatively, a multilevel or hierarchical linear model can be used to appropriately model this dependency. The following section will discuss hierarchical linear models.

HIERARCHICAL LINEAR MODEL

The conventional HLM models the dependency found in purely clustered data. Purely clustered data are those in which each level one unit (e.g., student) is a member of only one level two unit (e.g., school) although multiple level one units are members of each level two unit.

Unconditional HLM

Raudenbush and Bryk (2002) introduced a levels formulation for presenting HLM. Using this levels formulation, the two-level unconditional HLM, at level one is:

$$Y_{ij} = \beta_{0j} + e_{ij} \tag{2}$$

where Y_{ij} is the outcome for level one unit *i* clustered in level two unit *j*; β_{0j} is the average outcome for level two unit *j*; and e_{ij} is the level one residual which is assumed to be normally distributed with a mean of zero and a constant variance of σ^2. The level two model is presented as:

$$\beta_{0j} = \gamma_{00} + u_{0j} \tag{3}$$

where γ_{00} is the average outcome and u_{0j} is the level two residual which is assumed to be normally distributed with a mean of zero and a constant variance of τ_{00}. Equations 2 and 3 can be combined into a single equation:

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}. \tag{4}$$

Equation 4 shows the decomposition of the variability in the outcome into the part attributable to the level two unit j, u_{0j}, and the part attributable to the level one unit i, e_{ij}.

When estimating an HLM, unconditional models are typically estimated first in order to calculate the intra-class correlation coefficient (ICC). The ICC is the proportion of the total variance that is between level two units and is represented as follows (see, for example, Raudenbush & Bryk, 2002):

$$ICC = \frac{\tau_{00}}{\sigma^2 + \tau_{00}}. \tag{5}$$

The fully unconditional model (Equation 4) allows for estimation of the ICC (Equation 5). If the ICC is non-zero, the addition of predictors at each level can help explain some of this variability. For example, student and school characteristics can be used as predictors at level one and level two, respectively.

Conditional HLM

Predictors can be added at both level one (e.g., student level) and level two (e.g., school level) in a conditional HLM. For example, at level one, various student characteristics could be added as predictors and at level two, various school characteristics could be added. The model below includes one level one, X, and one level two, Z, predictor. At level one, Equation 2 would become:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij} \tag{6}$$

and the level two unconditional HLM (Equation 3) would become:

$$\begin{cases} \beta_{0j} = \gamma_{00} + \gamma_{01} Z_{.j} + u_{0j} \\ \beta_{1j} = \gamma_{10} \end{cases} \tag{7}$$

where γ_{00} is the average outcome when X_{ij} and $Z_{.j}$ are zero; γ_{10} is the change in the outcome for a one unit change in X_{ij}, holding all else constant; and γ_{01} is the change in the intercept when $Z_{.j}$ changes by one unit, holding all else constant. Equations 6 and 7 can be combined into the following single-equation model:

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_{.j} + u_{0j} + e_{ij}. \tag{8}$$

HLM allows for these predictors to be modeled as fixed or randomly varying across level two units. The intercept term, β_{0j}, in Equation 7 has been modeled as randomly varying across level two units while the slope, β_{1j}, has been modeled as fixed. Suppose that X represents a descriptor of student socioeconomic status (SES). If the researcher believes this variable's influence on the outcome is the same across schools, it would be modeled as fixed (see Equation 7). If, however, the researcher believes the relationship between X and Y changes depending on the school the student attends, then the coefficient, β_{1j}, could be modeled as randomly varying. Equation 7 would then be modified as follows:

$$\begin{cases} \beta_{0j} = \gamma_{00} + \gamma_{01} Z_{.j} + u_{0j} \\ \beta_{1j} = \gamma_{10} + u_{1j} \end{cases}. \tag{9}$$

In Equation 9, the u_{1j} error term was added indicating that β_{1j} varies randomly across level two units. Given the addition of the extra level two residual, u_{1j}, the distribution of the two level two residuals must be specified. The level two residuals (u_{0j} and u_{1j}) are typically assumed normally distributed with a mean vector of $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and covariance matrix of $\begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix}$. Additional combinations of predictors and assumptions about their pattern of relationships across level two units can be specified (see, for example, Raudenbush & Bryk, 2002).

Centering

Predictor variables can be used in their natural metric or centered. The location of the level one predictor variables gives the meaning to the level one intercept. For example, in Equation 6, β_{0j} is the expected outcome when X_{ij} is zero. In some cases, however, a value of zero on X_{ij} is not meaningful. For example, the Graduate Management Admission Test (GMAT) is scored on a scale from 200 to 800. If an individual's GPA is the outcome and their GMAT score is X, the intercept, β_{0j}, would then be the expected GPA when an individual scores a zero on the GMAT, which is not feasible. In cases such as this, a researcher may want to change the location of X_{ij} so the intercept will be meaningful. Two different centering techniques will be discussed here, namely, grand and group mean centering.

When grand mean centering is employed, the intercept is the expected value of the outcome for an individual whose X_{ij} value is equal to the grand mean, $\overline{X}_{..}$. Equation 6 would be modified as such:

$$Y_{ij} = \beta_{0j} + \beta_{1j}\left(X_{ij} - \overline{X}_{..}\right) + e_{ij} \tag{10}$$

Using the example presented above, the intercept would now be the expected GPA for an individual who scored at the average across all students. Enders and Tofighi (2007) recommend grand mean centering when the researcher wants to control for level one covariates, but the level two predictor is of substantive interest or the researcher is interested in the interactions among level two predictors.

When group mean centering is used, the intercept is the expected value of the outcome for an individual whose X_{ij} value is at their group mean, $\overline{X}_{.j}$. If group mean centering is employed, Equation 6 would become:

$$Y_{ij} = \beta_{0j} + \beta_{1j}\left(X_{ij} - \overline{X}_{.j}\right) + e_{ij} \tag{11}$$

In the GPA example, the intercept would now be the expected GPA for a student who scores at their school (if school is the level two unit) average. Group mean centering is recommended if one is primarily interested in the association between X (a level one predictor) and Y or if the interaction between two level one predictors is of interest (Enders & Tofighi, 2007).

Thus far, a data structure that has one level two unit per level one unit has been presented. However, there are other possible clustered data structures including, and of interest here, the multiple membership data structure which will be discussed in the next section.

MULTIPLE MEMBERSHIP DATA STRUCTURE

HLM can only be used when the dataset entails pure clusters, however, purely clustered data structures are not the only possibility in real world datasets. Purely clustered data are encountered when each level one unit (e.g., student) is a member of only one level two unit (e.g., school) with multiple level one units per level two unit. An example of such a dataset is depicted in Figure 1.

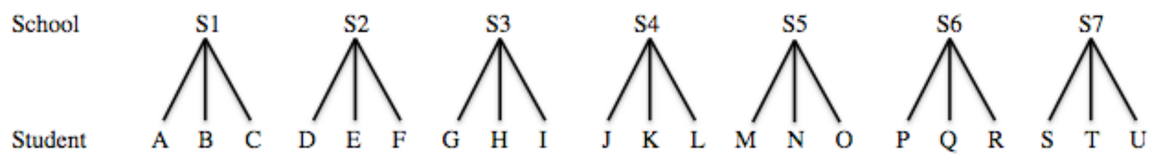

Figure 1. Purely Clustered Data

Students, however, change schools for a number of reasons such as a parental job transfer as a result of a promotion, one or both parents work in the military, or their parents are migrant laborers. The assumption when using HLM to model clustering of multiple students within each school is that each student is a member of only one school. If a student changes schools, then the student is no longer a member of just one school rather is a member of multiple schools. Such students are considered “mobile students”. Figure 2 is designed to clarify the distinction between a multiple membership data structure from a purely hierarchical data structure (as depicted in Figure 1).

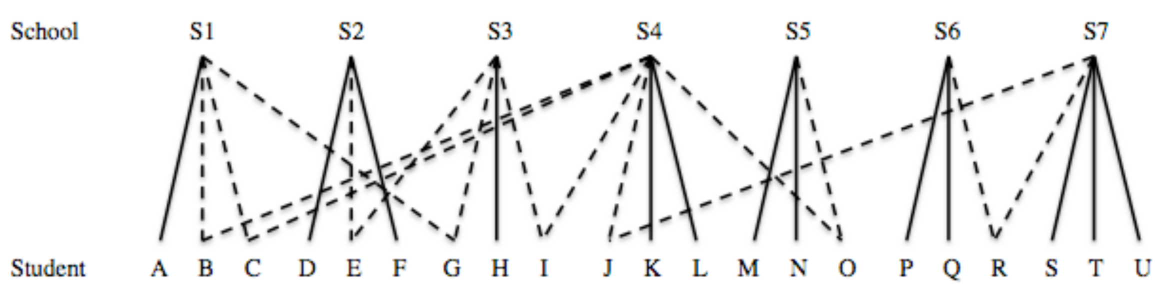

Figure 2. Multiple Membership Data Structure

In Figure 2, some students have been deemed mobile students and are depicted as attending more than one school. Solid lines indicate the student attended one school while dotted lines are reserved for mobile students and are connected to each school they attended (Beretvas, 2010). For example, student A (non-mobile) is shown to have attended only school one while student B is a mobile student and attended schools one and four.

The multiple membership data structure (see Figure 2) is not limited to the clustering of students within schools. For example, mobile residents have been members of multiple neighborhoods. In the medical field, patients can be members of (clustered in) one or more therapy groups. Given this study's focus is in the context of educational research, student mobility will be considered as the relevant source of a multiple membership data structure. The next section will discuss student mobility in general.

Student Mobility

Student mobility is a reality. There are two fundamental types of mobility including "normal" mobility involving a student's moving from, for example, an elementary to a middle school and "non-normative" mobility which involves, for example, a student's changing elementary school while still an elementary school student. Non-normative mobility is of interest in this study and from here on when mobility is discussed, the mobility refers to non-normative mobility.

Mantzicopoulos and Knutson (2000) found that, on average, students attended 1.30 schools with a standard deviation of 1.28 over a three year period (kindergarten through second grade). Over a four-year period (grades two through five), Gruman, Harachi, Abbott, and Catalano (2008) found that students changed schools 0.74 times with a standard deviation of 0.90.

As indicated above, many times students are members of multiple schools as a result of mobility. Use of conventional HLM requires that each level one unit be associated with only one level two unit and thus cannot handle the multiple membership data structures which result from this student mobility. However, the multiple membership random effects model (MMREM) is designed specifically for use with multiple membership data. The next section will cover MMREMs.

MULTIPLE MEMBERSHIP RANDOM EFFECTS MODEL

The MMREM is used to handle multiple membership data (i.e., data in which at least one level one unit is a member of multiple level two units). The parameterization of the unconditional and conditional models (Beretvas, 2010; Goldstein, 2010; Rasbash & Browne, 2001) will be discussed in this section.

Unconditional MMREM

The parameterization of the two-level unconditional MMREM at level one is as follows:

$$Y_{i\{j\}} = \beta_{0\{j\}} + e_{i\{j\}} \tag{12}$$

where $Y_{i\{j\}}$ is the outcome for level one unit (e.g., student) *i* who is a member of the set, $\{j\}$, of level two units (e.g., schools); $\beta_{0\{j\}}$ is the average outcome for the set of level two units, $\{j\}$; and $e_{i\{j\}}$ is the level one residual which is assumed normally distributed with a mean of zero and a constant variance of σ^2. The level two model is:

$$\beta_{0\{j\}} = \gamma_{00} + \sum_{h \in \{j\}} w_{ih} u_{0h} \tag{13}$$

where γ_{00} is the average outcome, w_{ih} is the weight associated with level one unit's association with level two unit h, and u_{0h} is the level two residual for level two unit h which is assumed normally distributed with a mean of zero and a constant variance of τ_{00}. Equations 12 and 13 can be combined to form the following equation:

$$Y_{i\{j\}} = \gamma_{00} + \sum_{h \in \{j\}} w_{ih} u_{0h} + e_{i\{j\}}. \tag{14}$$

There are two main distinctions between conventional HLM and MMREM. One difference is that instead of only one level two unit, j, per level one unit, i, as is the case with HLM (see Equation 4), MMREM allows for multiple level two units, $\{j\}$, to be associated with a single level one unit, i. Additionally, in comparing Equations 3 and 13, it is noted that weights are used in MMREM. These weights will be discussed in the next section.

Weights

Weights are assigned to each level two unit, j, included in the multiple membership dataset. A number of algorithms can be used to assign weights with the restriction that the weights must sum to one (Goldstein, 2010). Weights can be assigned equally or unequally. When assigning equal weights, each level two unit is assumed to have an equal contribution to the outcome. If, however, one level two unit might be hypothesized to have more of a contribution than the other(s), unequal weights could be assigned. For example, as shown in Figure 2, student B attended schools one and four. In a scenario in which these schools were equally weighted, Equation 13 would be as follows for that student:

$$\beta_{0\{1,4\}} = \gamma_{00} + 0.5u_{01} + 0.5u_{04}. \tag{15}$$

It may, however, be hypothesized that school one has more of an impact on the student's outcome than did school four. For example, if a student attended school one for first through third grade and attended school four for fourth grade, it may be hypothesized that school one has 75% (3/4) impact while school four has 25% (1/4) impact. If this is the case, then the following pattern of weights might be used:

$$\beta_{0\{1,4\}} = \gamma_{00} + 0.75u_{01} + 0.25u_{04}. \tag{16}$$

While researchers do have a choice of the values for the pattern of weights they assign to each set of level two units, research has indicated that the choice of weights' values does not greatly impact parameter estimates (Wolff Smith & Beretvas, 2012). Thus far, the discussion of MMREM has been focused on the unconditional model. As with HLM, conditional models are primarily used and will be presented in the next section.

Conditional MMREM

Predictors are typically added to the unconditional MMREM to help explain variability at each level. If one level one and one level two predictor (*X* and *Z*, respectively) are added to the model, then the level one model will become as follows (modification of Equation 12):

$$Y_{i\{j\}} = \beta_{0\{j\}} + \beta_{1\{j\}}X_{i\{j\}} + e_{i\{j\}} \tag{17}$$

where $\beta_{1\{j\}}$ is the change in $Y_{i\{j\}}$ when $X_{i\{j\}}$ changes by one unit, holding all else constant. The level two model is then (modification of Equation 13):

$$\begin{cases} \beta_{0\{j\}} = \gamma_{00} + \gamma_{01} \sum_{h \in \{j\}} w_{ih} Z_{.h} + \sum_{h \in \{j\}} w_{ih} u_{0h} \\ \beta_{1\{j\}} = \gamma_{10} \end{cases} \tag{18}$$

where γ_{00} is the average outcome when $X_{i\{j\}}$ and the weighted average of the level two predictors are zero; γ_{01} is the change in $\beta_{0\{j\}}$ for a one unit change in $Z_{.h}$, holding all else constant; and γ_{10} is the change in the outcome for a one unit change in $X_{i\{j\}}$, holding all else constant. Note that the intercept is modeled, here, as randomly varying and the slope is modeled as fixed across level two units. The combined model is then:

$$Y_{i\{j\}} = \gamma_{00} + \gamma_{10} X_{i\{j\}} + \gamma_{01} \sum_{h \in \{j\}} w_{ih} Z_{.h} + \sum_{h \in \{j\}} w_{ih} u_{0h} + e_{i\{j\}} \,. \tag{19}$$

Additional or fewer predictors could be modeled as well as different patterns of fixed and randomly varying intercept and slopes (see, for example, Beretvas, 2010; Goldstein, 2010).

Note that when using conventional HLM, the level two predictor is simply $Z_{.j}$ (see Equation 8). However, the level two predictor in the MMREM is: $\sum_{h \in \{j\}} w_{ih} Z_{.h}$. In the case of MMREM, the level two predictor is a weighted average of the level two predictor's values across the set of schools attended by the mobile student. For example, student B in Figure 2 attended schools one and four. Using equal weights, Equation 19 would be modified as shown below for this student:

$$Y_{B\{1,4\}} = \gamma_{00} + \gamma_{10} X_{B\{1,4\}} + \gamma_{01}\left(0.5 Z_{.1} + 0.5 Z_{.4}\right) + \left(0.5 u_{01} + 0.5 u_{04}\right) + e_{B\{1,4\}} \tag{20}$$

If, however, conventional HLM were to be used instead of MMREM, the researcher would be required to choose which school would be identified as the level two unit. If the last school attended is used as the level two unit, Equation 8 would be modified as shown below for student B:

$$Y_{B4} = \gamma_{00} + \gamma_{10} X_{B4} + \gamma_{01} Z_{.4} + u_{04} + e_{B4} \tag{21}$$

The level two predictor is not a weighted average of the two schools' predictor values attended nor is the level two error a weighted composite of the two schools' errors as is the case when using MMREM (see Equation 20).

The next section will provide a review of the methodological research performed on the MMREM.

Methodological Work With the MMREM

Some methodological research on the use and estimation of the MMREM has been conducted. First, researchers compared use of conventional HLM versus MMREM when a multiple membership data structure was present. When using conventional HLM, the multiple membership data structure was ignored and only the last school that a mobile student attended was modeled. In a school effectiveness study, Goldstein, Burgess, and McConnell (2007) found that when the multiple membership data structure was ignored and conventional HLM was used, the contribution of the school (level two unit) to the outcome was underestimated. Leckie (2009) built on the research by Goldstein et al. (2007). In addition to looking at the clustering of students within schools, the clustering in neighborhoods was examined. Similar to the results found by Goldstein et al. (2007), Leckie (2009) found that MMREM corrects for the underestimation of the school and neighborhood variances that would otherwise be present when using conventional HLM.

In the medical field, Chandola, Clarke, Wiggins, and Bartley (2005) investigated physical and mental health functioning. In this paper, the authors explored a multiple membership data structure in which individuals were clustered within households and residential areas. Three models were estimated, namely, (1) individuals (level one) clustered in households (level two), (2) individuals (level one) clustered in residential areas (level two), and (3) individuals (level one) clustered in households (level two) clustered in residential areas (level three). For each model and each outcome (physical and mental health functioning), an unconditional and conditional HLM and MMREM were estimated. Model 1 resulted in a slightly higher household variance when MMREM was used as compared with the conventional HLM. The results of Model 2 differed across outcomes and model type (unconditional versus conditional). When the mental health functioning score was the outcome, MMREM resulted in a reduction in residential area variance as compared with conventional HLM for both the unconditional and conditional models. When the physical functioning score was the outcome, the unconditional MMREM resulted in an increase in residential area variance. However, the conditional MMREM resulted in slightly smaller residential area variance across models estimated. When using MMREM, Model 3 resulted in a higher household variance and lower residential area variance as compared with conventional HLM. Overall, accounting for the multiple membership structure appeared to increase the estimates of the household variance while the residential area variance decreased when both household and area clustering were included in the model as well as other risk factors (the predictors in the conditional model). The authors concluded that longitudinal analysis should include any moves between household and/or residential area over time. Thus, MMREM should be used to model this type of data over conventional HLM.

Chung and Beretvas (2012) performed a simulation study that compared the use of conventional HLM and MMREM where only the last school attended was modeled when using conventional HLM. They found negative bias in the coefficient of the level two predictor and level two variance component estimates as well as overestimation of the level one variance component when the multiple membership data structure was ignored. Only a small amount of positive bias was found in the level two variance component estimates when MMREM was used. This is, however, to be expected as Markov Chain Monte Carlo (MCMC) estimation was used to estimate the models. MCMC is known to overestimate the highest level's variance components of the random effects (Browne & Draper, 2006). In an extension of this line of research, Wolff Smith and Beretvas (2011) compared three models for handling and assessing mobility, namely, HLM-*Delete*, HLM-*Last School*, and MMREM. When using the HLM-*Delete* model, students were deleted from the dataset if they were members of multiple schools. HLM-*Last School* modeled only the last school attended. Finally, the MMREM modeled all schools each student attended. The authors found that the level two predictor's coefficient had some substantial negative bias when the two HLM approaches were used. Additionally, the level one mobility predictor's coefficient was found to be substantially positively biased and the level two mobility predictor's coefficient was found to be substantially negatively biased when using the HLM-*Last School* approach. Substantial bias was also found in some conditions when using MMREM as well, but no consistent pattern was discerned. The level two variance component was found to be substantially negatively biased when using HLM-*Last School* and substantially positively biased when using HLM-*Delete* and MMREM. As discussed earlier, MCMC estimation is known to overestimate the variance component at the highest level (Browne & Draper, 2006), so the MMREM results for estimates of the higher level variance component were expected.

For the mobility pattern generated in Wolff Smith and Beretvas' (2011) study, results for the HLM-*Delete* procedure were found to closely mimic the MMREM results. In this paper and in the paper by Chung and Beretvas (2012), mobility was randomly assigned to students. As such, mobility was not modeled as a function of student characteristics such as SES. Thus, in the Wolff Smith and Beretvas (2011) paper, when mobile students' data were deleted from the analysis, a randomly selected group of student data was deleted. As a result, the HLM-*Delete* and MMREM procedures would be similar as the models are run on a similar population of students. However, mobility is not a randomly assigned condition. On the contrary, many correlates of mobility are also correlates of achievement. Thus, removing mobile students' data from an analysis should change the characteristics of the population whose data are being analyzed.

Applied Research on Mobility

This section will briefly summarize applied research on mobility, including sources of mobility and as well as how it has been found to relate to student achievement.

Factors Related to Mobility

Researchers have investigated a variety of factors that relate to mobility. For example, Hanushek, Kain, and Rivkin (2004) looked at the trends of mobility for fourth through seventh graders. They found that if the student had ever qualified for low income status (low SES), they changed schools more than students who never qualified as low SES. Additionally, this research has supported that a larger percentage of minority students are mobile as compared with Caucasians. Wright (1999) also found this to be the case and concluded that SES and minority status were equal influences on mobility. Been, Gould Ellen, Schwartz, and Weinstein (2011) investigated the effects of foreclosure on mobility. The authors found that the current foreclosure crisis has been associated with an increase in the incidence of mobility.

Family structure has also been found to influence student mobility. For example, Astone and Mclanahan (1994) found that children who are members of single parent and step-families are more likely to be mobile than children in two-parent families. Similarly, Adduci (1990) found that children living in a single parent household are more likely to be mobile than children living with two parents. Implications of this student mobility will be addressed in the following section.

Effects of Mobility

In 2004, Mehana and Reynolds performed a meta-analysis of studies involving mobility and academic achievement (math and reading) between 1975 and 1994 with a focus on elementary aged (kindergarten through sixth grade) students. Across the studies, they found an overall effect size for mobility of -0.22 for math achievement and -0.25 for reading achievement. The negative effect size indicates that mobile students had lower math and reading achievement scores than did non-mobile students.

More studies have been performed assessing the impact of mobility on student achievement since 1994. Many studies after 1994 have resulted in the same conclusions as were found in Mehana and Reynolds' (2004) meta-analysis, student mobility is related to lower achievement scores (e.g., Mantzicopoulos & Knutson, 2000; Strand & Demie, 2007; Temple & Reynolds, 1999). However, Strand and Demie (2006) and Tucker, Marx, and Long (1998) both concluded that student mobility does not have a negative impact on educational achievement scores. This was, however, only the case in the Tucker et al. (1998) paper when children were brought up in a household with both biological parents. If the child was a member of any other family structure, mobility was found to negatively impact school life.

While some of the findings dispute this, the majority of studies conclude that mobility does negatively affect students' achievement. This is not the case, however, for children in military families. Children of military families do not appear to be as negatively affected as children from civilian families (Cramer & Dorsey, 1970). This may be, in part, due to the support structures created by the military to make the moves easier for those involved.

Scenarios in Previous Methodological Research

Previous methodological simulation research on the MMREM neglected to model mobility as an endogenous predictor of achievement, rather mobility status was randomly assigned. It is far more likely that mobility is not randomly assigned. Instead, there are variables, represented in Figure 3 using the symbol *P*, that likely increase a student's propensity to be mobile that are also related to an outcome that mobility affects (for example, achievement). Some examples of *P* include family structure (e.g., single parent family), family income, and SES. In Figure 3, *Y* is the outcome (e.g., achievement score), *M* is mobility status, *X* is a student characteristic (e.g., pre-test score), *P* is a proxy for variables affecting a student's propensity to be mobile (e.g., combination of family structure, family income, SES) as well as *Y*, and *e* represents the residual for *Y*.

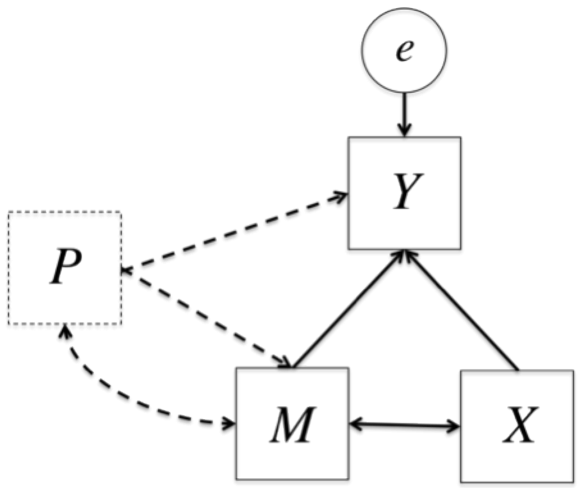

Figure 3. Depiction of Endogeneity

If a researcher is interested in assessing whether mobility predicts an outcome, *Y*, but neglects to also include the variable(s), *P*, (that is also related to *M* and *Y*) in the prediction of *Y* (as in Figure 4), then mobility will be an endogenous rather than exogenous predictor. This means that whether or not a student is mobile is correlated with *P*, which itself affects *Y*. Given the potential endogeneity of a student's likelihood of mobility, it is possible that this could result in biased parameter estimates due to the selection bias associated with the non-mobile students whose data are not removed. A model that might be estimated is shown in Figure 4.

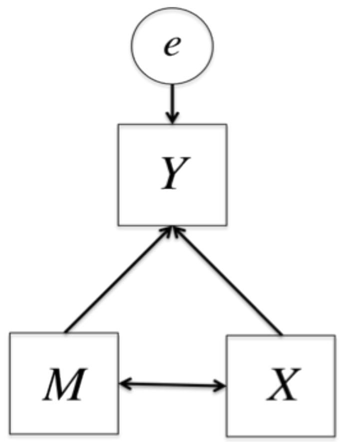

Figure 4. Potential Estimated Model

As shown in Figure 4, *P* has not been included, thus the endogeneity of *M* is not modeled. Bias results ensue because the model must compensate in some way for this missing term. Various econometrics textbooks have addressed the issue of "omitted variable bias" (e.g., Barreto & Howland, 2005; Greene, 1993). This is the bias associated with the removal of an independent variable that is both correlated with the dependent variable as well as one or more independent variables. In their econometrics book, Barreto and Howland (2005) discussed that in ordinary least squares regression, failure to include a variable in the model that is correlated with the dependent variable will result in biased and inconsistent estimates.

If mobility were really endogenous, it is expected the results of the Wolff Smith and Beretvas (2011) paper would be affected. It is anticipated that the relative parameter bias using HLM-*Delete* and MMREM will differ rather than being so similar as the mobile students who are removed in the HLM-*Delete* analysis constitute a different sub-population than the non-mobile students. Generating more realistic patterns of differences between mobile and non-mobile students should provide more authentic comparisons of model estimates that can better inform applied researchers. To make the generated dataset more realistic, this study will not randomly assign mobility, rather it will involve generating mobility status as a function of some student level characteristics that are unmodeled in the estimating models.

In addition, previous research was performed under the ideal scenario in which researchers have access to the full set of level two units associated with each level one unit. In reality, however, this may not be the case. Hill and Goldstein (1998) addressed this issue for data with two higher level units with missing identifiers for one of the pair. In their paper, Hill and Goldstein (1998) suggested a method for handling missing level two identifiers in the context of a two-level cross-classified dataset. In the example dataset used, the level one unit represented students while the level two classification variables were class and school (j_1 and j_2, respectively). The school attended by the student was known for every student, however, for some students the class affiliation was unknown or "missing". Hill and Goldstein (1998) suggested an ad hoc procedure for handling this pattern of missing identifiers. Under this procedure, a set of weights is calculated for each student with an unknown classification (here, class) membership. These weights represent the square root of the probability that the student is in each of the possible classes within the school attended by the student. For example, say student i, for whom class membership was unknown, attended school j_2. Thus, the weight formula for student i could be calculated using:

$$\pi_{ij_2} = \sqrt{\frac{1}{n_{j_2}}} \tag{22}$$

where π_{ij_2} represents the weight to use for student i who is known to have attended school j_2, although for whom it is not known in which class the student was enrolled, and n_{j_2} represents the number of classes in school j_2. Thus, for example, assume that school j_2 had nine classes. The weight associated with each class in school j_2 for student i would be calculated using Equation 22 to be $\sqrt{1/9} = 1/3$. While the weights will not sum to one, Hill and Goldstein (1998) did not discuss this and thus it does not appear to be problematic.

The Hill and Goldstein (1998) procedure was used on the Victorian Quality Schools Study, an educational effectiveness longitudinal study that took place over the period of three years. Just over 19% of the students in the study were missing a classroom identification. Hill and Goldstein (1998) compared three different models, namely, (1) a conditional model (achievement predicted by grade level) using class identification weights (see Equation 22) for students with a missing identifier, (2a) a conditional model (achievement predicted by grade level as well as four student level characteristics) using identification weights for those with a missing identifier, and (2b) Model 2a, but where individuals with a missing class identifier were removed listwise from the dataset. While the true parameter values are unknown for this real dataset, comparisons were made across models. The most notable findings resulted from the comparison between Models

2a and 2b. The fixed effects and random effects variance component parameter and standard error estimates were similar when students who had missing class identifiers were removed from the analysis as compared with the analysis based on Hill and Goldstein's (1998) ad hoc procedure for handling the missing identifiers. Given the lack of differences found between the results of Models 2a and 2b, it seems likely that, for this particular dataset, the likelihood of a student's class identifier being missing was unrelated to the student outcome. Regardless of the similar results found in this study, the authors point out that deleting students who are missing identifiers is not efficient and bias in the estimates of the fixed and random parts of the model would likely ensue. Thus, researchers should not delete students with missing identifiers, rather they should use a procedure such as the one presented in the article to account for the missing data.

Use of the ad hoc procedure for handling the missing class identifier proposed by Hill and Goldstein (1998) requires some knowledge of the relevant number of identifiers for the missing classes per additional higher level classification (school). When analyzing data with missing identifiers for a classification at the highest level in a data's structure (for example, when missing some level two identifiers in a two-level MMREM), the Hill and Goldstein (1998) procedure cannot be used. Thus, another ad hoc procedure, termed the "MMREM-*Unique*" procedure, was derived for the current study as an alternative for use when level two identifiers are missing in a two-level MMREM. Discussion of the MMREM-*Unique* procedure will follow the discussion of the HLM-*Delete* and MMREM-*Delete* procedures commonly used by researchers. Note that the term "procedure" will be used to refer to each combination of model and procedure for ease of expression.

HLM-*Delete* Procedure

When using conventional HLM, each level one unit can be a member of only one level two unit. As such, a multiple membership data structure cannot be modeled. The HLM-*Delete* procedure will involve removing all mobile students from the analysis and the resulting reduced dataset will be analyzed. It is important to include this procedure as it provides an authentic match to what is used by many applied researchers and thus provides an important comparison for this research.

MMREM-*Delete* Procedure

When employing the MMREM-*Delete* procedure, the level one unit (for example, student) will be deleted from the analysis if they are missing at least one level two unit (here, school) identifier. The MMREM will then be estimated using the resulting reduced dataset. Clearly, employing this technique will decrease the sample size of the dataset being analyzed and will result in a decrease of statistical power. In addition to this negative result, biased parameter estimates could ensue as a result of the removal of mobile students who might constitute a different population than non-mobile students.

MMREM-*Unique* Procedure

The MMREM-*Unique* procedure will involve assigning a placeholder identifier that is unique for each case that is missing an identifier. This equates to assigning the student to a fictional school at the time point where the identifier is missing. Given there will be a single placeholder school identifier for each student who is missing a real identifier, estimation of that fictional school's effect should not greatly impact estimation of real schools' effects. It is not expected that the MMREM-*Unique* procedure will result in unbiased parameter estimates as it will involve a misspecification. However, it is expected that parameter recovery using the MMREM-*Unique* procedure will be superior to that of the HLM-*Delete* and MMREM-*Delete* procedures. While all three procedures will involve a misspecification, the reduced dataset makes it seem likely that MMREM-*Delete* and HLM-*Delete* will not recover the parameters as well as MMREM-*Unique*. An example will be presented in the next section to show how these three procedures work.

EXAMPLE OF AD HOC PROCEDURES FOR HANDLING MISSING IDENTIFIERS

An example can be used to demonstrate these three ad hoc procedures (HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*) for handling missing level two identifiers. Note that this example is for illustration purposes only. The example dataset used for this illustration is shown in Table 1. Note that the data are the same as that depicted in Figure 2 but with a few schools identified as missing. Missing school identifiers are represented using a ".". Each procedure discussed previously will handle the missing school identifiers differently. Table 1 shows the results of using the HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique* procedures for handling missing identifiers.

Table 1

Comparison of Ad Hoc Procedures

			Original		HLM *Delete*		MMREM *Delete*		MMREM *Unique*	
ID	*m*	*mis*	S1	S2	S1	S2	S1	S2	S1	S2
A	0	0	1	1	1	1	1	1	1	1
B	1	1	.	4	--	--	--	--	1000	4
C	1	0	1	4	--	--	1	4	1	4
D	0	0	2	2	2	2	2	2	2	2
E	1	1	.	3	--	--	--	--	1001	3
F	0	0	2	2	2	2	2	2	2	2
G	1	1	.	1	--	--	--	--	1002	1
H	0	0	3	3	3	3	3	3	3	3
I	1	0	3	4	--	--	3	4	3	4
J	1	0	4	7	--	--	4	7	4	7
K	0	0	4	4	4	4	4	4	4	4
L	0	0	4	4	4	4	4	4	4	4
M	0	0	5	5	5	5	5	5	5	5
N	0	0	5	5	5	5	5	5	5	5
O	1	1	.	5	--	--	--	--	1003	5
P	0	0	6	6	6	6	6	6	6	6
Q	0	0	6	6	6	6	6	6	6	6
R	1	1	.	7	--	--	--	--	1004	7
S	0	0	7	7	7	7	7	7	7	7
T	0	0	7	7	7	7	7	7	7	7
U	0	0	7	7	7	7	7	7	7	7

Note. *m* = mobility status (1 = mobile), *mis* = missing a school identifier (1 = missing a school identifier), S = school, "." = missing data, "--" = record not included.

As seen in Table 1, the dataset and school identifiers differ across the three ad hoc procedures (HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*). Using the HLM-*Delete* procedure, all students who are mobile are deleted from the dataset and using the MMREM-*Delete* procedure, all students with a missing school identifier are deleted from the dataset. For the data appearing in Table 1, eight students' data would be removed using HLM-*Delete* and five students' data would be removed using MMREM-*Delete*, reducing the dataset from 21 to 13 and 16 cases, respectively. Note that school identifiers for some mobile students (for example, student J) are not missing and these cases are not removed from the analysis when MMREM-*Delete* is employed but are removed when HLM-*Delete* is employed. When using either *delete* procedure, the mobility rate of the dataset changes. With all 21 students included, the mobility rate is 38.1% (8 of 21 cases). This rate decreases to 18.8% (3 of 16 cases) when the students with missing school identifiers are removed from the analysis (MMREM-*Delete*) and decreases to 0% when HLM-*Delete* is used. The mobility rate will always decrease using the *delete* procedures as only mobile students will be removed.

When using the MMREM-*Unique* procedure, the missing school identifiers are replaced with a *unique* (pseudo-) school identifier and the weights are adjusted as if the student really attended the placeholder school. For example, student B was missing an identifier at school one. As such, a unique school identifier, 1000, that is not shared with any other student, was assigned as their school one identifier.

STATEMENT OF PURPOSE

Given the reality of students' mobility, procedures are necessary for handling missing school identifiers. The three procedures presented (HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*) seem to provide reasonable alternatives although it is unclear how their use might affect resulting parameter estimates. The current study is intended to inform applied researchers and practitioners working in schools and school districts about the optimal analytic techniques and procedures necessary for handling a specific real-world data complication. This study will compare the effects of using three procedures for handling missing school identifiers (HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*). Additionally, in generating the data, mobility will not be assigned randomly as has been done in past simulation research. Instead, the endogeneity of mobility will be generated although not modeled in the estimating models. The comparison between the three ad hoc procedures (HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*) will entail a simulation study that will assess parameter recovery and standard error estimation for both fixed effects and random effects variance components. Manipulated conditions will include the intra-class correlation coefficient, the percent of mobile students, the number of level two units (here, schools), and the percent of mobile students with a missing identifier. The details of the simulation study method will be presented in the next chapter.

Chapter 3: Method

The current simulation study investigated the implications of using three procedures, namely HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*, to handle missing level two (here, school) identifiers. In addition, mobility was not randomly assigned as in previous simulation research, rather the endogeneity of mobility was generated but not modeled in the estimating models. This simulation study was used to assess the parameter and standard error bias of parameters estimated after employing these three procedures.

GENERATING CONDITIONS

Various conditions were manipulated in this simulation study to investigate their impact on parameter recovery. The manipulated conditions were the percent mobility (10%, 20%), the intra-class correlation coefficient (ICC) (15%, 25%), the number of level two units (50, 100), and the percent of mobile level one units (here, student) with missing level two units (here, school) (25%, 50%). Table 2 shows the combination of design conditions that were explored in this study.

Table 2

Combination of Design Conditions

Manipulated Conditions			
Mobility	ICC	Level Two Units	Mobile Level One Units with Missing Level Two Identifiers
10%	15%	50	25%
10%	15%	50	50%
10%	15%	100	25%
10%	15%	100	50%
10%	25%	50	25%
10%	25%	50	50%
10%	25%	100	25%
10%	25%	100	50%
20%	15%	50	25%
20%	15%	50	50%
20%	15%	100	25%
20%	15%	100	50%
20%	25%	50	25%
20%	25%	50	50%
20%	25%	100	25%
20%	25%	100	50%

Mobility Rate

In their investigation of large-scale national longitudinal datasets, Chung and Beretvas (2012) found small and moderate mobility rates to be 10% and 20%, respectively. As such, approximately the same mobility percentages were used in this simulation study.

Intra-class Correlation Coefficient

The residual intra-class correlation coefficient (ICC) is defined as the proportion of the total variance that is attributed to level two after covariates are entered into the model (see Equation 5). Spybrook and Raudenbush (2009) discussed four different studies that analyzed ICCs for educational outcomes. In the majority of the studies, they found ICC values ranging from 0.10 to 0.25. Based on these findings, generating values of 0.15 and 0.25 (15% and 25%, respectively) were used for the ICCs examined in this study. In the Wolff Smith and Beretvas (2011) study, the value of the level one variance was generated to be 200 across conditions. The current study used the same value for the level one variance. Therefore, the value of the generated level two variance was calculated as follows:

$$\tau_{00} = \frac{200(ICC)}{1 - ICC} \tag{23}$$

(see Equation 5).

Number of Level Two Units Assigned to Each Level One Unit

Gruman et al. (2008) found that over a four year period, students (level one units) change schools an average of 0.74 times with a standard deviation of 0.90. Similarly, over a three year period, Mantzicopoulos and Knutson (2000) found that students attended an average of 1.30 schools with a standard deviation of 1.28. Therefore, as a starting point for an investigation of procedures for handling missing level two identifiers, this simulation study allowed mobile students to attend a maximum of two schools.

Level Two Units

Chung and Beretvas (2012) performed a simulation study assessing estimation of a two-level MMREM in which the authors manipulated the number of level two units. They estimated the MMREM with 50 and 100 level two units and found that most of the parameters were reasonably well recovered with 50 level two units. Because this paper's focus is on the missingness of level two units, it was deemed appropriate to use both 50 and 100 to assess whether the number of level two units impacts the recovery of parameters under each ad hoc procedure.

Missing Identifiers

This study allowed approximately 25% and 50% of the mobile students to be lacking a school identifier. The 50% condition was used to show what happens in what is thought to be the worst-case scenario. This means that, of the simulated dataset, the following percent of students were both mobile and missing a school identifier: 2.5% (10% mobility, 25% missing condition), 5% (10% mobility, 50% missing condition and 20% mobility, 25% missing condition), and 10% (20% mobility, 50% missing condition). As a first step in this type of investigation, if a student is missing a school identifier, it will occur at the first time point. Future research should be designed that investigates the impact of missing identifiers at a variety of time points. In the next section, the generating and estimating models will be discussed.

Generation and Estimation

One thousand datasets per combination of conditions were generated using MLwiN software (version 2.24, 2011). Details of the model generation and estimation are provided in the following sections.

Generating and Estimating Models

The generating model included three student level predictors. Future research can extend such a model to include school level predictors as well. The student level predictors of interest in this study are a group mean centered, continuously scored student level predictor (e.g., pre-test score); a group mean centered, dichotomously scored mobility predictor; and a group mean centered, continuously scored predictor that represents the positive opportunity for a student to be mobile (see Figure 3). The process for generating the predictors is described in a section below. The generating model is as follows:

$$Y_{i\{j\}} = \gamma_{00} + \gamma_{10}\left(X_{i\{j\}} - \overline{X}_{.\{j\}}\right) + \gamma_{20}\left(M_{i\{j\}} - \overline{M}_{.\{j\}}\right) + \gamma_{30}\left(P_{i\{j\}} - \overline{P}_{.\{j\}}\right) + \sum_{h\in\{j\}} w_{ih}u_{0h} + e_{i\{j\}} \quad (24)$$

where Y is the outcome variable, X is a student level predictor, M is the student level mobility predictor, and P is a proxy for positive opportunity to be mobile (e.g., combination of family structure, family income, SES). Following the suggestion of Enders and Tofighi (2007), the level one predictors have been group mean centered since the interest is in the association between level one predictors and the outcome. While a group mean centered, dichotomous predictor may seem unorthodox, Enders and Tofighi (2007) explained this can and should be done under certain circumstances. The resulting intercept of a model where the only predictor is a group mean centered, binary variable is interpreted as the unadjusted mean for that particular group (here, school). For convenience, the group mean each predictor variable is centered around was the school at the last time point. It is assumed the characteristics of schools across time (for mobile students) remain reasonably similar. Future research should assess alternative procedures for handling group mean centering with multiple membership structures.

It seems likely that researchers would not include every variable associated with mobility in the model, thus resulting in endogeneity bias. To mimic real world scenarios, the estimating model did not include *P* as this variable is what is considered here as the unobserved source of endogeneity in mobility that is related to *M*, *X*, and *Y* (see Figure 4). The MMREM estimating model is as follows:

$$Y_{i\{j\}} = \gamma_{00} + \gamma_{10}\left(X_{i\{j\}} - \overline{X}_{.\{j\}}\right) + \gamma_{20}\left(M_{i\{j\}} - \overline{M}_{.\{j\}}\right) + \sum_{h\in\{j\}} w_{ih}u_{0h} + e_{i\{j\}} \tag{25}$$

and the HLM estimating model is as follows:

$$Y_{ij} = \gamma_{00} + \gamma_{10}\left(X_{ij} - \overline{X}_{.j}\right) + u_{0j} + e_{ij}. \tag{26}$$

Fixed Effects

Five fixed effects are present in the generating model (see Equation 24), namely, γ_{00}, γ_{10}, γ_{20}, and γ_{30} which represent the intercept and effects of the mobility predictor, student level predictor, and positive opportunity variable, respectively. Generating values for γ_{00}, γ_{10}, γ_{20}, and γ_{30} were set to 100, 1, -0.5, and 1 mimicking the values used in the Wolff Smith and Beretvas (2011) study. Wolff Smith and Beretvas (2011) did not use a predictor for positive opportunity, but it is believed such a predictor would have around the same effect as the student level predictor; therefore, the generating value of γ_{30} was set to one as well.

Random Effects

The variance of the level one residual, $e_{i\{j\}}$, was fixed across conditions at 200 following previous research (Wolff Smith & Beretvas, 2011). Therefore, each student's level one residual was randomly selected from a normal distribution with a mean of zero and a variance of 200. The variance of the level two residual (τ_{00}) depended on the condition (see Equation 23).

Weights

As discussed previously, MMREM weights can be assigned equally or unequally (see, for examples, Equations 15 and 16). Wolff Smith and Beretvas (2012) found the assignment of weights does not greatly impact the parameter estimates. As such, for simplification purposes, equal weights were assigned. Thus, for a student that attended one school, a weight of one was assigned to that school. For a student that attended two schools, weights of 1/2 were assigned to each school the student attended.

Predictors

The predictors in this study (M, X, and P; see Equation 24) were generated as related to one another. Corr (1982) found that mobility and SES were correlated with $r = -0.04$, mobility and math achievement with $r = -0.05$, and SES and math achievement with $r = 0.26$. Similarly, White (1982) found the correlation between SES and academic achievement to be 0.22. In their meta-analysis, Mehana and Reynolds (2004) found the overall correlations between student mobility and reading achievement and between student mobility and math achievement to be -0.31 and -0.17, respectively. X is a student level predictor such as math or reading achievement. P is a proxy for positive opportunity for mobility and can be thought of as a combination of several characteristics including SES. Due to these similarities between the variables in the current study and the studies presented, the correlations above could be used to generate the three related variables in the generating model (see Equation 24). However, it is expected that the values of these correlations will be attenuated when dichotomizing the mobility variable. As such, the values observed were increased to counteract this effect. Thus, the total correlations that were used to generate the standard normal predictor variables were $r_{M,X} = -0.2$, $r_{M,P} = -0.2$, and $r_{X,P} = 0.4$.

When identifying the mobile students and mobile students with a missing identifier, a certain level of error must be present. As such, the procedure described below was used to generate M, X, and P such that error is introduced in the generation of the mobile students and mobile students with a missing identifier. Raudenbush, Bryk, Cheong, and Congdon (2004) describe a model in which proportional odds is assumed and is as follows:

$$\eta_{mij} = \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} + \sum_{m=2}^{M} \delta_m . \tag{27}$$

Under this model it is assumed that each person falls into one category, m, and there are a total of M categories. In the present study, there are three possible categories, namely, non-mobile, mobile, and mobile with a missing identifier. Categories $m - 1$ and m are separated by δ_m. Thus, for $M = 3$, the thresholds can be modeled as:

$$
\begin{aligned}
\eta_{1ij} &= \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} \\
\eta_{2ij} &= \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} + \delta_2
\end{aligned}
. \qquad (28)
$$

X and P are the predictors of mobility and mobile students with a missing identifier. Therefore, the log odds ratio of non-mobile versus mobile is:

$$
\log\left[\frac{p(\text{mobility}_i)}{1 - p(\text{mobility}_i)}\right] = \beta_0 + \beta_1 X_i + \beta_2 P_i \qquad (29)
$$

and the log odds ratio of mobile versus mobile with a missing identifier is:

$$
\log\left[\frac{p(\text{missing}_i)}{1 - p(\text{missing}_i)}\right] = \beta_0 + \beta_1 X_i + \beta_2 P_i + \delta_2 . \qquad (30)
$$

Equations 29 and 30 can be converted to probabilities using simple algebra. The probability equations are as follows:

$$
p(\text{mobility}_i) = \frac{e^{\beta_0 + \beta_1 X_i + \beta_2 P_i}}{1 + e^{\beta_0 + \beta_1 X_i + \beta_2 P_i}} \qquad (31)
$$

and

$$
p(\text{missing}_i) = \frac{e^{\beta_0 + \beta_1 X_i + \beta_2 P_i + \delta_2}}{1 + e^{\beta_0 + \beta_1 X_i + \beta_2 P_i + \delta_2}} . \qquad (32)
$$

Note that these probabilities are cumulative.

A brief simulation was run to find reasonable values for β_0, β_1, β_2, and δ_2 (see Equations 31 and 32) such that $r_{M,X} \approx -0.2$, $r_{M,P} \approx -0.2$, and $r_{X,P} \approx 0.4$ and the mobility and missing identifier rates approximately match the condition (see Table 2). Table 3 shows the values that seem reasonable for β_0, β_1, β_2, and δ_2 by condition. Note that each sample of 50,000 students produced slightly different values for the parameters and the values shown below are from one replication.

Table 3

Parameter Values for Mobility and Missing Identifier Probability Equations

	Condition			
	10% *M* 2.5% *MM*	10% *M* 5% *MM*	20% *M* 5% *MM*	20% *M* 10% *MM*
$\beta_{0,M}$	2.40	2.40	1.50	1.50
$\beta_{0,MM}$	3.90	3.20	3.20	2.40
δ_2	1.50	0.80	1.70	0.90
β_1	0.45	0.45	0.45	0.45
β_2	0.45	0.45	0.45	0.45
%*M*	10.11%	10.05%	20.57%	20.78%
%*MM*	2.59%	5.00%	4.93%	10.09%
$r_{M,X}$	-0.178	-0.183	-0.237	-0.229
$r_{M,P}$	-0.182	-0.184	-0.239	-0.229
$r_{X,P}$	0.407	0.400	0.400	0.396

Note. *M* = mobile, *MM* = mobile with a missing identifier.

The following steps were followed to generate *M*, *X*, and *P*. First, β_0, β_1, β_2, and δ_2 were set based on the condition (see Table 3). *X* and *P* were then randomly selected from a bivariate normal distribution with a mean of zero and a standard deviation of one. Using

Equations 31 and 32, the probabilities of mobility and missing were calculated. A random number was then generated from a uniform [0,1] distribution to compare with the current probabilities. If the random number selected from the uniform [0,1] distribution was less than or equal to the probability of mobility, then the student was identified as non-mobile. In the event that the random number selected from the uniform [0,1] distribution was between the probability of mobility and probability of missing, then the student was deemed mobile. Finally, if the random number selected from the uniform [0,1] distribution was greater than or equal to the probability of missing, the student was identified as a mobile student with a missing identifier. Note that this procedure will represent a worst-case scenario for student mobility in that students with lower *X* and *P* values are more likely to be mobile and mobile with a missing identifier.

The student level predictor, *X*, was converted to a normally distributed variable with a mean of 50 and a standard deviation of 10 following previous research (Chung & Beretvas, 2012; Wolff Smith and Beretvas, 2011). The proxy variable, *P*, for positive opportunity to be mobile was converted to the same scale. In the next section, the generation of school characteristics will be discussed.

Generation of Schools and Their Characteristics

The number of schools generated was based on the condition (50 or 100). Students were first assigned probabilities of attending a lower performing school. Mobility that is negatively related to achievement is being generated in this study, thus students who are not mobile have a lower probability of attending a lower performing (in terms of achievement) school (here, 0.3) than a mobile student or mobile student missing an identifier (here, 0.7). To introduce some error, this probability was compared against a number that was randomly selected from a uniform [0,1] distribution. If the random number selected from a uniform [0,1] distribution was less than or equal to the probability of attending a lower performing school, then the student was assigned to a lower performing school. Otherwise, they were assigned to a higher performing school.

Out of the 50 or 100 schools (condition dependent), a certain number of schools were deemed lower performing. The goal was to have approximately the same number of students per school. As such, a brief simulation was performed with 50,000 students to determine how many schools would need to be labeled “low performing” to ensure that the number of students per school was approximately the same. For the 10% and 20% mobility conditions, it was found that approximately 34% and 38%, respectively, of the students attended a low performing school. As such, for the 10% mobility conditions, 17 and 34 schools were deemed “low performing” for the 50 and 100 school conditions, respectively. For the 20% mobility conditions, 19 and 38 schools were labeled as “low performing” for the 50 and 100 school conditions, respectively.

As discussed earlier, if a student was mobile, they attended two schools. Also, if a student was mobile and missing an identifier, the missing school identifier was generated to occur at the first time point. To begin, all students were assigned to their first school. For mobile students, a second school was assigned. To assign this school, one was added to the identification number for school one. If, however, the student attended the highest number school at the first time point, they were assigned to the lowest number school at the second time point. This matches the procedure used in previous simulation research (e.g., Meyers & Beretvas, 2006). If a student were missing a school identifier, they would not have attended one of the 50 or 100 schools as those are within the study's scope. Therefore, each of those students was assigned to a school outside of the study's focal area. It is reasonable to assume that each student moving into the focal area of the study is from a different school. While not inconceivable, it is not very likely that a group of students move together from the same unidentified school. The pattern of data that was generated then closely matched what a *unique* dataset looks like. While this may advantage the performance of the *unique* procedure over the *delete* procedures, this pattern seems to provide a reasonably sound match with reality. The schools were assigned as shown in Tables 4 and 5.

Table 4

School Assignment for the 10% Mobility Conditions

Number of Level 2 Units	Low/High Performing Schools	Mobility status	Attended Low/High Performing Schools	S1	S2
50	L: 1-17 H: 18-50	*NM*	Low performing	Randomly selected between 1 and 17	Same as S1
			High performing	Randomly selected between 18 and 50	Same as S1
		M	Low performing	Randomly selected between 1 and 17	If S1 = 17, then S2 = 1; otherwise S1 + 1
			High performing	Randomly selected between 18 and 50	If S1 = 50, then S2 = 18; otherwise S1 + 1
		MM	Low performing	Unique identifier	Randomly selected between 1 and 17
			High performing	Unique identifier	Randomly selected between 18 and 50
100	L: 1-34 H: 35-100	*NM*	Low performing	Randomly selected between 1 and 34	Same as S1
			High performing	Randomly selected between 35 and 100	Same as S1
		M	Low performing	Randomly selected between 1 and 34	If S1 = 34, then S2 = 1; otherwise S1 + 1
			High performing	Randomly selected between 35 and 100	If S1 = 100, then S2 = 35; otherwise S1 + 1
		MM	Low performing	Unique identifier	Randomly selected between 1 and 34
			High performing	Unique identifier	Randomly selected between 35 and 100

Note. L = low performing school, H = high performing school, *NM* = non-mobile, *M* = mobile, *MM* = mobile with a missing identifier, S1 = school 1, S2 = school 2.

Table 5

School Assignment for the 20% Mobility Conditions

Number of Level 2 Units	Low/High Performing Schools	Mobility status	Attended Low/High Performing Schools	S1	S2
50	L: 1-19 H: 20-50	*NM*	Low performing	Randomly selected between 1 and 19	Same as S1
			High performing	Randomly selected between 20 and 50	Same as S1
		M	Low performing	Randomly selected between 1 and 19	If S1 = 19, then S2 = 1; otherwise S1 + 1
			High performing	Randomly selected between 20 and 50	If S1 = 50, then S2 = 20; otherwise S1 + 1
		MM	Low performing	Unique identifier	Randomly selected between 1 and 19
			High performing	Unique identifier	Randomly selected between 20 and 50
100	L: 1-38 H: 39-100	*NM*	Low performing	Randomly selected between 1 and 38	Same as S1
			High performing	Randomly selected between 39 and 100	Same as S1
		M	Low performing	Randomly selected between 1 and 38	If S1 = 38, then S2 = 1; otherwise S1 + 1
			High performing	Randomly selected between 39 and 100	If S1 = 100, then S2 = 39; otherwise S1 + 1
		MM	Low performing	Unique identifier	Randomly selected between 1 and 38
			High performing	Unique identifier	Randomly selected between 39 and 100

Note. L = low performing school, H = high performing school, *NM* = non-mobile, *M* = mobile, *MM* = mobile with a missing identifier, S1 = school 1, S2 = school 2.

Residuals were assigned based on whether the school was identified as low or high performing. The residuals for high performing schools were randomly selected from a normal distribution with a mean of 0.5 and a standard deviation of 1 while the low performing school residuals were randomly selected from a normal distribution with a mean of -0.5 and a standard deviation of 1. Once these residuals were generated, they were transformed such that the residual distribution was approximately normal with a mean of zero and variance of the condition-specific τ_{00} value (see Equation 23).

Estimation Procedure

MLwiN software's (version 2.24, 2011) Markov Chain Monte Carlo (MCMC) estimation procedure which employs both Gibbs and Metropolis-Hastings sampling was used to estimate the parameters for each of the three procedures investigated (HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*). Browne (2009) provides additional details about the estimation procedure. For each set of conditions one chain was run with 50,000 iterations and a burn-in of 5,000 matching what Chung and Beretvas (2012) found to be sufficient for stable estimation of a reasonably simple MMREM model.

ANALYSES

Estimates of the γ_{00}, γ_{10}, γ_{20}, σ^2, and τ_{00} parameters in the estimation model (see Equations 25 and 26) were compared in terms of both their relative parameter and standard error bias (Hoogland & Boomsma, 1998). Details of the bias measures are provided in the next section.

Relative Parameter Bias

Relative parameter bias was calculated for each of the parameter estimates, namely, γ_{00}, γ_{10}, γ_{20}, σ^2, and τ_{00}. The relative parameter bias was calculated using the formula below:

$$B(\hat{\theta}_j) = \frac{\bar{\hat{\theta}}_j - \theta_j}{\theta_j} \tag{33}$$

(Hoogland & Boomsma, 1998) where θ_j is the true value of parameter j and $\bar{\hat{\theta}}_j$ is the estimate of parameter j averaged across the 1,000 replications of each condition. According to Hoogland and Boomsma (1998), a parameter is considered substantially biased if the relative parameter bias is larger than a magnitude of 0.05. This means that a parameter is considered substantially biased if it differs from the true value of the parameter by over 5%.

Relative Standard Error Bias

Similar to the relative parameter bias, relative standard error bias was calculated for each of the parameter estimates. The following formula was used to calculate the relative standard error bias:

$$B\left(s_{\hat{\theta}_j}\right) = \frac{\bar{\hat{s}}_{\hat{\theta}_j} - \sigma_{\hat{\theta}_j}}{\sigma_{\hat{\theta}_j}} \tag{34}$$

(Hoogland & Boomsma, 1998) where $\bar{\hat{s}}_{\hat{\theta}_j}$ is the average standard error estimate of parameter j across the 1,000 replications and $\sigma_{\hat{\theta}_j}$ is the true standard error of parameter j. The standard error of the distribution of the estimated j (1,000 estimates) was considered

$\sigma_{\hat{\theta}_j}$ here. A parameter is considered to have substantial standard error bias when the relative standard error bias exceeds 0.10 in magnitude (Hoogland & Boomsma, 1998). This means that the standard error of the estimate differs from the true standard error by over 10%.

Analysis of Variance

Analyses of variance (ANOVAs) were performed to explore whether certain condition specifications (e.g., mobility rate, ICC value) affected the relative parameter and standard error bias estimates. ANOVAs were conducted for each procedure (MMREM-*Unique*, MMREM-*Delete*, and HLM-*Delete*) separately with the relative bias (either parameter or standard error) as the dependent variable and the condition specifications as the independent variables. Main effects and two-way interactions were included in all ANOVAs. Given the sample size, practical rather than statistical significance was interpreted. Cohen (1977) deemed partial eta squared $\left(\eta_p^2\right)$ values of 0.01, 0.06, and 0.14 as small, moderate, and large effect sizes, respectively. Here, a conservative cutoff was used in that an effect was considered practically significant if the associated partial eta squared value was 0.01 or greater (e.g., Kirk, 1995).

Chapter 4: Results

This chapter presents the results of the current simulation study which compared three different procedures for handling missing school identifiers, namely, HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*. When using HLM-*Delete*, all mobile students (including those who were and were not missing school identifiers) were removed from the analysis. All students with missing school identifiers were removed from the analysis when MMREM-*Delete* was employed. Finally, with MMREM-*Unique* a unique school identifier was used to replace the missing identifier. As such, MMREM-*Unique* was the only procedure that included the entire dataset. All estimates were obtained using MCMC estimation in MLwiN (version 2.24, 2011) with 50,000 iterations and a burn-in of 5,000 iterations. Relative parameter and standard error biases were summarized for each condition (see Table 2).

Relative Parameter Bias

The relative parameter bias was calculated for the estimates of the intercept (γ_{00}); coefficient of the level one predictor, X (γ_{10}); coefficient of the level one mobility predictor, M (γ_{20}); level one variance component (σ^2); and level two variance component (τ_{00}). The parameters were considered substantially biased if the relative parameter bias was greater than a magnitude of 0.05 (Hoogland & Boomsma, 1998).

Intercept, γ_{00}

The relative parameter bias of the estimates of the intercept, γ_{00}, by condition and procedure estimated is presented in Table 6. As indicated in Table 6, no substantial relative parameter bias was detected in any condition.

Table 6

Relative Parameter Bias of the Intercept, γ_{00}, Estimates

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	0.006	0.010	0.019
10%	15%	50	50%	0.009	0.016	0.022
10%	15%	100	25%	0.006	0.010	0.017
10%	15%	100	50%	0.011	0.017	0.022
10%	25%	50	25%	0.008	0.012	0.022
10%	25%	50	50%	0.011	0.019	0.025
10%	25%	100	25%	0.010	0.014	0.021
10%	25%	100	50%	0.015	0.022	0.027
20%	15%	50	25%	0.007	0.014	0.033
20%	15%	50	50%	0.010	0.024	0.037
20%	15%	100	25%	0.009	0.016	0.031
20%	15%	100	50%	0.013	0.025	0.036
20%	25%	50	25%	0.009	0.018	0.037
20%	25%	50	50%	0.011	0.027	0.040
20%	25%	100	25%	0.013	0.021	0.036
20%	25%	100	50%	0.018	0.033	0.042

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

X Coefficient, γ_{10}

Table 7 contains the relative parameter bias for estimates of the coefficient of *X*, γ_{10}. As shown in Table 7, substantial positive relative parameter bias was found across all conditions and procedures. The bias was the same across procedures to the second decimal place. The values ranged from 0.363 to 0.384 indicating that the coefficient of X, γ_{10}, was overestimated by between 36.3% and 38.4%. The ANOVA results for all three procedures indicated that only the percent mobility condition, m, affected the severity of the bias [MMREM-*Unique*: $F(1, 15989) = 297.58$, $p < 0.001$, $\eta_p^2 = 0.018$; MMREM-*Delete*: $F(1, 15989) = 365.31$, $p < 0.001$, $\eta_p^2 = 0.022$; HLM-*Delete*: $F(1, 15989) = 223.18$, $p < 0.001$, $\eta_p^2 = 0.014$]. The bias was consistently lower in conditions with a higher mobility percentage, m, (MMREM-*Unique*: $M_{m=10\%} = 0.380$, $M_{m=20\%} = 0.367$; MMREM-*Delete*: $M_{m=10\%} = 0.380$, $M_{m=20\%} = 0.364$; HLM-*Delete*: $M_{m=10\%} = 0.381$, $M_{m=20\%} = 0.368$).

Table 7

Relative Parameter Bias of the X Coefficient, γ_{10}, Estimates

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	***0.376***	***0.376***	***0.376***
10%	15%	50	50%	***0.381***	***0.381***	***0.381***
10%	15%	100	25%	***0.382***	***0.382***	***0.384***
10%	15%	100	50%	***0.381***	***0.381***	***0.382***
10%	25%	50	25%	***0.382***	***0.381***	***0.381***
10%	25%	50	50%	***0.377***	***0.377***	***0.378***
10%	25%	100	25%	***0.380***	***0.380***	***0.381***
10%	25%	100	50%	***0.380***	***0.380***	***0.382***
20%	15%	50	25%	***0.365***	***0.364***	***0.368***
20%	15%	50	50%	***0.368***	***0.366***	***0.368***
20%	15%	100	25%	***0.366***	***0.363***	***0.366***
20%	15%	100	50%	***0.366***	***0.364***	***0.367***
20%	25%	50	25%	***0.367***	***0.365***	***0.369***
20%	25%	50	50%	***0.367***	***0.365***	***0.369***
20%	25%	100	25%	***0.367***	***0.365***	***0.369***
20%	25%	100	50%	***0.367***	***0.364***	***0.368***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Mobility Coefficient, γ_{20}

The relative parameter bias for the estimates of the mobility coefficient, *M*, γ_{20}, is shown in Table 8. Note that this parameter was not estimated for HLM-*Delete* as only non-mobile students' data were analyzed. Very substantial positive relative parameter bias was found across all procedures that estimated this parameter (MMREM-*Delete* and MMREM-*Unique*). Across all conditions, MMREM-*Delete* had less bias than MMREM-*Unique* ($M_{\text{MMREM-}Delete}$ = 6.865, $M_{\text{MMREM-}Unique}$ = 7.984). ANOVA results indicate that no condition had a substantial impact on the mobility coefficient, γ_{20}, estimate in either procedure.

Table 8

Relative Parameter Bias of the Mobility Coefficient, γ_{20}, Estimates

Condition				Procedure	
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*
10%	15%	50	25%	***7.586***	***7.011***
10%	15%	50	50%	***7.877***	***6.935***
10%	15%	100	25%	***7.799***	***7.026***
10%	15%	100	50%	***8.453***	***7.137***
10%	25%	50	25%	***7.954***	***7.267***
10%	25%	50	50%	***8.370***	***7.080***
10%	25%	100	25%	***8.230***	***7.236***
10%	25%	100	50%	***8.697***	***6.675***
20%	15%	50	25%	***7.498***	***6.857***
20%	15%	50	50%	***7.804***	***6.563***
20%	15%	100	25%	***7.658***	***6.823***
20%	15%	100	50%	***8.106***	***6.546***
20%	25%	50	25%	***7.565***	***6.896***
20%	25%	50	50%	***7.813***	***6.545***
20%	25%	100	25%	***7.900***	***6.887***
20%	25%	100	50%	***8.426***	***6.361***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Level One Variance, σ^2

The average relative parameter bias of the level one variance component, σ^2, estimates is shown in Table 9. Substantial positive relative parameter bias was found across conditions and procedures. To two decimal places, the bias was the same across all three procedures ($M_{\text{MMREM-}Unique}$ = 0.414, $M_{\text{MMREM-}Delete}$ = 0.414, $M_{\text{HLM-}Delete}$ = 0.416) and ranged from 0.409 to 0.423 indicating that the level one variance, σ^2, was overestimated by between 40.9% and 42.3%. The ANOVA results indicated that no condition had a significant effect on the bias.

Table 9

Relative Parameter Bias of the Level One Variance Component, σ^2, Estimates

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	***0.420***	***0.420***	***0.423***
10%	15%	50	50%	***0.416***	***0.416***	***0.418***
10%	15%	100	25%	***0.415***	***0.415***	***0.416***
10%	15%	100	50%	***0.414***	***0.414***	***0.415***
10%	25%	50	25%	***0.414***	***0.414***	***0.416***
10%	25%	50	50%	***0.416***	***0.416***	***0.417***
10%	25%	100	25%	***0.412***	***0.412***	***0.413***
10%	25%	100	50%	***0.415***	***0.415***	***0.415***
20%	15%	50	25%	***0.414***	***0.414***	***0.420***
20%	15%	50	50%	***0.415***	***0.417***	***0.420***
20%	15%	100	25%	***0.410***	***0.409***	***0.411***
20%	15%	100	50%	***0.411***	***0.412***	***0.414***
20%	25%	50	25%	***0.413***	***0.413***	***0.417***
20%	25%	50	50%	***0.412***	***0.412***	***0.416***
20%	25%	100	25%	***0.410***	***0.409***	***0.411***
20%	25%	100	50%	***0.414***	***0.412***	***0.415***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Level Two Variance, τ_{00}

The resulting relative parameter bias for the level two variance component, τ_{00}, estimates is presented in Table 10.

Table 10

Relative Parameter Bias of the Level Two Variance Component, τ_{00}, Estimates

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	***-0.113***	***-0.135***	***-0.181***
10%	15%	50	50%	***-0.113***	***-0.157***	***-0.186***
10%	15%	100	25%	***-0.129***	***-0.150***	***-0.199***
10%	15%	100	50%	***-0.148***	***-0.188***	***-0.219***
10%	25%	50	25%	-0.045	***-0.067***	***-0.113***
10%	25%	50	50%	-0.036	***-0.076***	***-0.104***
10%	25%	100	25%	***-0.091***	***-0.109***	***-0.148***
10%	25%	100	50%	***-0.109***	***-0.142***	***-0.166***
20%	15%	50	25%	***-0.087***	***-0.120***	***-0.184***
20%	15%	50	50%	***-0.079***	***-0.142***	***-0.171***
20%	15%	100	25%	***-0.122***	***-0.151***	***-0.213***
20%	15%	100	50%	***-0.141***	***-0.195***	***-0.231***
20%	25%	50	25%	-0.023	***-0.054***	***-0.123***
20%	25%	50	50%	0.003	***-0.057***	***-0.101***
20%	25%	100	25%	***-0.093***	***-0.118***	***-0.171***
20%	25%	100	50%	***-0.096***	***-0.145***	***-0.182***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Substantial negative relative parameter bias was found across all conditions for both the MMREM-*Delete* and HLM-*Delete* estimates of the level two variance component, τ_{00}, and was found across the majority of the conditions for MMREM-

Unique (12 out of the 16 conditions). MMREM-*Unique* was found to consistently have the least amount of bias in comparison to the other two procedures (MMREM-*Delete* and HLM-*Delete*) across conditions while HLM-*Delete* was found to have the largest amount of bias ($M_{\text{MMREM-}\textit{Unique}}$ = -0.089, $M_{\text{MMREM-}\textit{Delete}}$ = -0.125, $M_{\text{HLM-}\textit{Delete}}$ = -0.168). The magnitude of bias ranged from 0.3% to 14.8% for MMREM-*Unique*, 5.4% to 19.5% for MMREM-*Delete*, and 10.1% to 23.1% for HLM-*Delete*.

Substantial negative relative parameter bias was found in the MMREM-*Unique* estimates when the ICC was 15% and when, together, the ICC was 25% and the number of level two units (c) was 100. No bias was found when, together, ICC = 25% and c = 50. Substantial negative bias was found across all conditions when the MMREM-*Delete* and HLM-*Delete* procedures were employed.

The ICC was found to significantly affect the relative parameter bias across all three procedures [MMREM-*Unique*: $F(1, 15989) = 352.29$, $p < 0.001$, $\eta_p^2 = 0.022$; MMREM-*Delete*: $F(1, 15989) = 397.74$, $p < 0.001$, $\eta_p^2 = 0.024$; HLM-*Delete*: $F(1, 15989) = 377.74$, $p < 0.001$, $\eta_p^2 = 0.023$]. As the ICC increased (15% to 25%), the magnitude of the relative parameter bias of the level two variance component, τ_{00}, estimates decreased (MMREM-*Unique*: $M_{\text{ICC=15\%}}$ = -0.117, $M_{\text{ICC=25\%}}$ = -0.061; MMREM-*Delete*: $M_{\text{ICC=15\%}}$ = -0.155, $M_{\text{ICC=25\%}}$ = -0.096; HLM-*Delete*: $M_{\text{ICC=15\%}}$ = -0.198, $M_{\text{ICC=25\%}}$ = -0.139). The number of level two units, c, also affected the bias significantly [MMREM-*Unique*: $F(1, 15989) = 345.57$, $p < 0.001$, $\eta_p^2 = 0.021$; MMREM-*Delete*: $F(1, 15989) = 269.53$, $p < 0.001$, $\eta_p^2 = 0.017$; HLM-*Delete*: $F(1, 15989) = 225.77$, $p < 0.001$, $\eta_p^2 =$

0.014]. The relative parameter bias of the level two variance component, τ_{00}, estimates was found to increase in magnitude as the number of level two units, c, increased (MMREM-*Unique*: $M_{c=50}$ = -0.061, $M_{c=100}$ = -0.116; MMREM-*Delete*: $M_{c=50}$ = -0.101, $M_{c=100}$ = -0.150; HLM-*Delete*: $M_{c=50}$ = -0.145, $M_{c=100}$ = -0.191).

RELATIVE STANDARD ERROR BIAS

Relative standard error (*SE*) bias was calculated for each parameter estimated in each procedure (HLM-*Delete*, MMREM-*Delete*, and MMREM-*Unique*) and set of conditions. Relative standard error bias values above a magnitude of 0.10 were deemed substantially biased following the recommendation by Hoogland and Boomsma (1998).

Intercept, γ_{00}

The relative *SE* bias of the estimates of the intercept, γ_{00}, by condition and procedure is shown in Table 11. Substantial positive relative *SE* bias was found across all procedures and conditions. Of the three procedures investigated and across all conditions, HLM-*Delete* resulted in the smallest amount of bias ($M_{\text{MMREM-}Unique}$ = 0.392, $M_{\text{MMREM-}Delete}$ = 0.333, $M_{\text{HLM-}Delete}$ = 0.256). The bias for MMREM-*Unique* ranged from 24.0% to 66.3%, MMREM-*Delete* ranged from 14.8% to 57.3%, and HLM-*Delete* ranged from 11.3% to 47.0%. The positive bias indicates that the *SE* of the intercept was consistently overestimated.

Table 11

Relative Standard Error Bias of the Intercept, γ_{00}, Estimates

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	***0.391***	***0.377***	***0.282***
10%	15%	50	50%	***0.327***	***0.241***	***0.204***
10%	15%	100	25%	***0.460***	***0.445***	***0.317***
10%	15%	100	50%	***0.296***	***0.277***	***0.208***
10%	25%	50	25%	***0.663***	***0.534***	***0.461***
10%	25%	50	50%	***0.435***	***0.377***	***0.328***
10%	25%	100	25%	***0.562***	***0.573***	***0.470***
10%	25%	100	50%	***0.378***	***0.341***	***0.274***
20%	15%	50	25%	***0.373***	***0.260***	***0.149***
20%	15%	50	50%	***0.294***	***0.148***	***0.113***
20%	15%	100	25%	***0.379***	***0.303***	***0.192***
20%	15%	100	50%	***0.283***	***0.207***	***0.195***
20%	25%	50	25%	***0.461***	***0.413***	***0.266***
20%	25%	50	50%	***0.307***	***0.250***	***0.193***
20%	25%	100	25%	***0.429***	***0.379***	***0.301***
20%	25%	100	50%	***0.240***	***0.204***	***0.146***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

The ANOVAs resulted in three main effects and two interactions being significantly related to the relative *SE* bias of the intercept, γ_{00}, estimates across all three procedures. The first of these common significant main effects was the mobility percentage, m, [MMREM-*Unique*: $F(1, 15989) = 3396.7$, $p < 0.001$, $\eta_p^2 = 0.175$; MMREM-*Delete*: $F(1, 15989) = 6102.0$, $p < 0.001$, $\eta_p^2 = 0.276$; HLM-*Delete*: $F(1, 15989) = 6414.6$, $p < 0.001$, $\eta_p^2 = 0.286$]. As the percent of mobile students, m, in the dataset increased from 10% to 20%, the relative *SE* bias decreased (MMREM-*Unique*: $M_{m=10\%} = 0.439$, $M_{m=20\%} = 0.346$; MMREM-*Delete*: $M_{m=10\%} = 0.396$, $M_{m=20\%} = 0.271$; HLM-*Delete*: $M_{m=10\%} = 0.318$, $M_{m=20\%} = 0.194$). The amount of decrease was the largest for HLM-*Delete*. It was also found that the true ICC value was significantly related to the relative *SE* bias [MMREM-*Unique*: $F(1, 15989) = 2777.8$, $p < 0.001$, $\eta_p^2 = 0.148$; MMREM-*Delete*: $F(1, 15989) = 4033.7$, $p < 0.001$, $\eta_p^2 = 0.201$; HLM-*Delete*: $F(1, 15989) = 3976.7$, $p < 0.001$, $\eta_p^2 = 0.199$] such that as the ICC increased (15% to 25%), the bias increased as well (MMREM-*Unique*: $M_{\mathrm{ICC}=15\%} = 0.350$, $M_{\mathrm{ICC}=25\%} = 0.434$; MMREM-*Delete*: $M_{\mathrm{ICC}=15\%} = 0.282$, $M_{\mathrm{ICC}=25\%} = 0.384$; HLM-*Delete*: $M_{\mathrm{ICC}=15\%} = 0.207$, $M_{\mathrm{ICC}=25\%} = 0.305$). The amount of increase in the bias was found to be similar across procedures. The final main effect that had a substantial effect on the bias was the percent of mobile students missing an identifier [MMREM-*Unique*: $F(1, 15989) = 8212.4$, $p < 0.001$, $\eta_p^2 = 0.339$; MMREM-*Delete*: $F(1, 15989) = 9357.1$, $p < 0.001$, $\eta_p^2 = 0.369$; HLM-*Delete*: $F(1, 15989) = 3957.1$, $p < 0.001$, $\eta_p^2 = 0.198$]. Across all procedures, as the percent of those mobile students missing an identifier, *mis*, increased, the relative *SE* bias decreased (MMREM-*Unique*: $M_{mis=25\%} = 0.465$, $M_{mis=50\%} = 0.320$; MMREM-*Delete*: $M_{mis=25\%} = 0.411$, $M_{mis=50\%} = 0.256$; HLM-*Delete*: $M_{mis=25\%} = 0.305$, $M_{mis=50\%} = 0.208$). While significant main effects of the mobility percentage, ICC, and percent of mobile students with a missing identifier were common across all procedures, MMREM-*Unique* also had a significant effect of the number of level two units [$F(1, 15989) = 309.5, p < 0.001, \eta_p^2 = 0.019$]. As the number of level two units, c, increased, the bias decreased significantly ($M_{c=50} = 0.406, M_{c=100} = 0.378$).

These main effects, however, do depend on another condition evidenced by the significant two-way interactions found. Two two-way interactions were significant across all procedures, namely the ICC by number of level two units [MMREM-*Unique*: $F(1, 15989) = 512.3, p < 0.001, \eta_p^2 = 0.031$; MMREM-*Delete*: $F(1, 15989) = 492.9, p < 0.001, \eta_p^2 = 0.030$; HLM-*Delete*: $F(1, 15989) = 321.2$, $p < 0.001$, $\eta_p^2 = 0.020$] and ICC by percent of mobile students with missing identifiers [MMREM-*Unique*: $F(1, 15989) = 760.9$, $p < 0.001$, $\eta_p^2 = 0.045$; MMREM-*Delete*: $F(1, 15989) = 282.2$, $p < 0.001$, $\eta_p^2 = 0.017$; HLM-*Delete*: $F(1, 15989) = 742.2$, $p < 0.001$, $\eta_p^2 = 0.044$]. In conditions where the ICC was 15%, more positive bias resulted when $c = 100$ as opposed to $c = 50$ (MMREM-*Unique*: $M_{\text{ICC}=15\%,\ c=50} = 0.346$, $M_{\text{ICC}=15\%,\ c=100} = 0.354$; MMREM-*Delete*: $M_{\text{ICC}=15\%,\ c=50} = 0.257$, $M_{\text{ICC}=15\%,\ c=100} = 0.308$; HLM-*Delete*: $M_{\text{ICC}=15\%,\ c=50} = 0.187$, $M_{\text{ICC}=15\%,\ c=100} = 0.228$) whereas the opposite trend appeared when the ICC was 25% (MMREM-*Unique*: $M_{\text{ICC}=25\%,\ c=50} = 0.466$, $M_{\text{ICC}=25\%,\ c=100} = 0.402$; MMREM-*Delete*: $M_{\text{ICC}=25\%,\ c=50} = 0.394$, $M_{\text{ICC}=25\%,\ c=100} = 0.374$; HLM-*Delete*: $M_{\text{ICC}=25\%,\ c=50} = 0.312$, $M_{\text{ICC}=25\%,\ c=100} = 0.298$). Figure 5 depicts this interaction for the MMREM-*Unique* estimates, while Figures 6 and 7 display patterns found with MMREM-*Delete* and HLM-*Delete* mean estimates, respectively. In the scenarios where the ICC was 15%, there was a larger decrease in relative *SE* bias between 25% and 50% of mobile students missing an identifier, *mis*, (MMREM-*Unique*: $M_{\text{ICC=15\%},\ mis=25\%} = 0.401$, $M_{\text{ICC=15\%},\ mis\ =\ 25\%} = 0.300$; MMREM-*Delete*: $M_{\text{ICC=15\%},\ mis=25\%} = 0.346$, $M_{\text{ICC=15\%},\ mis\ =\ 25\%} = 0.218$; HLM-*Delete*: $M_{\text{ICC=15\%},\ mis=25\%} = 0.235$, $M_{\text{ICC=15\%},\ mis\ =\ 25\%} = 0.180$) than there was in the ICC = 25% scenarios (MMREM-*Unique*: $M_{\text{ICC=25\%},\ mis=25\%} = 0.529$, $M_{\text{ICC=25\%},\ mis\ =\ 25\%} = 0.340$; MMREM-*Delete*: $M_{\text{ICC=25\%},\ mis=25\%} = 0.475$, $M_{\text{ICC=25\%},\ mis\ =\ 25\%} = 0.293$; HLM-*Delete*: $M_{\text{ICC=25\%},\ mis=25\%} = 0.375$, $M_{\text{ICC=25\%},\ mis\ =\ 25\%} = 0.235$). Figures 8 through 10 depict this relationship for the MMREM-*Unique*, MMREM-*Delete*, and HLM-*Delete*, respectively.

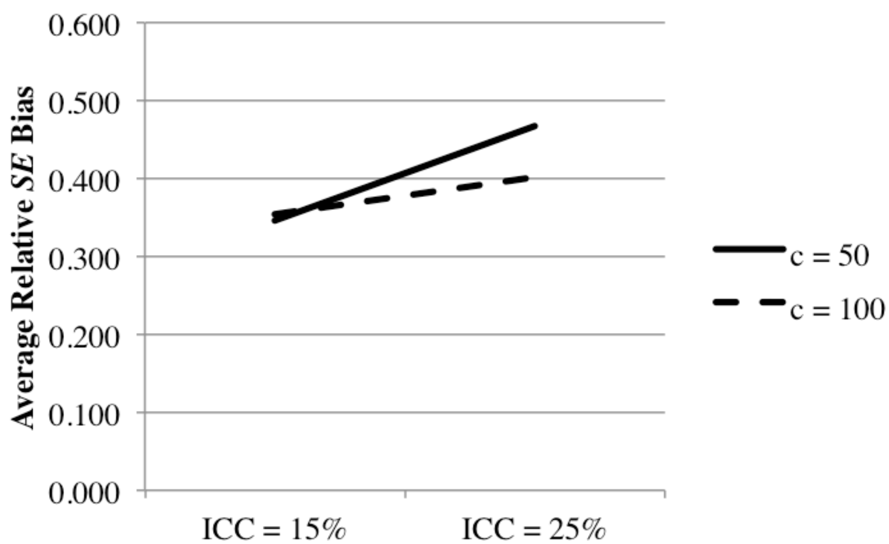

Figure 5. ICC and Level Two Units Interaction Effect on the Relative *SE* Bias of γ_{00} Estimates for MMREM-*Unique*

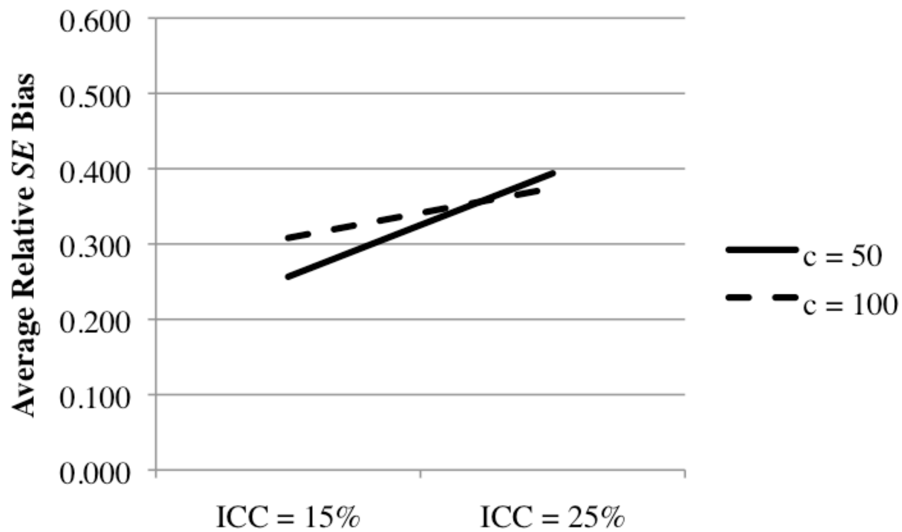

Figure 6. ICC and Number of Level Two Units Interaction Effect on the Relative *SE* Bias of γ_{00} Estimates for MMREM-*Delete*

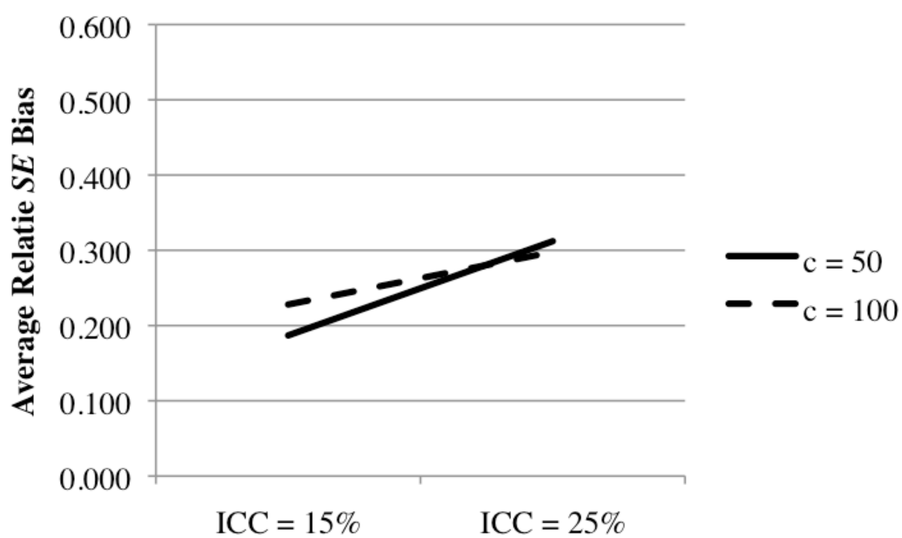

Figure 7. ICC and Number of Level Two Units Interaction Effect on the Relative *SE* Bias of γ_{00} Estimates for HLM-*Delete*

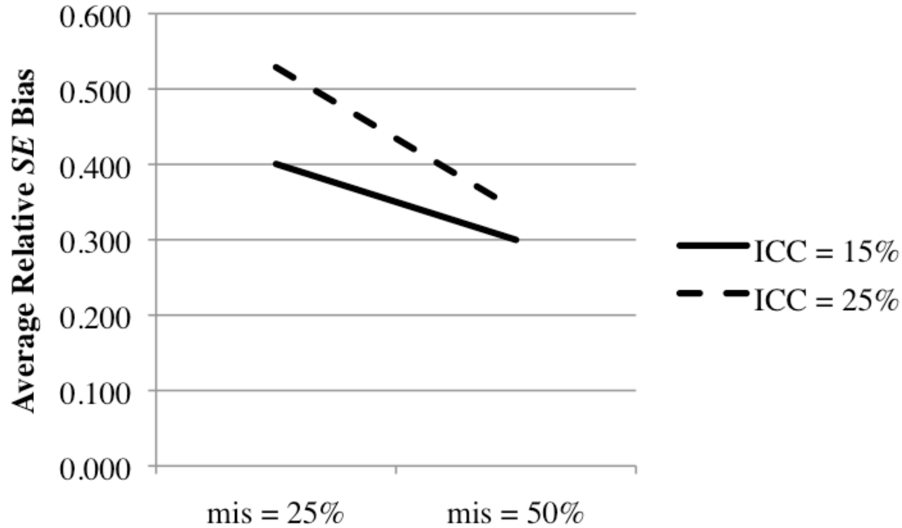

Figure 8. ICC and Percent of Mobile Students Missing an Identifier Interaction Effect on the Relative *SE* Bias of γ_{00} Estimates for MMREM-*Unique*

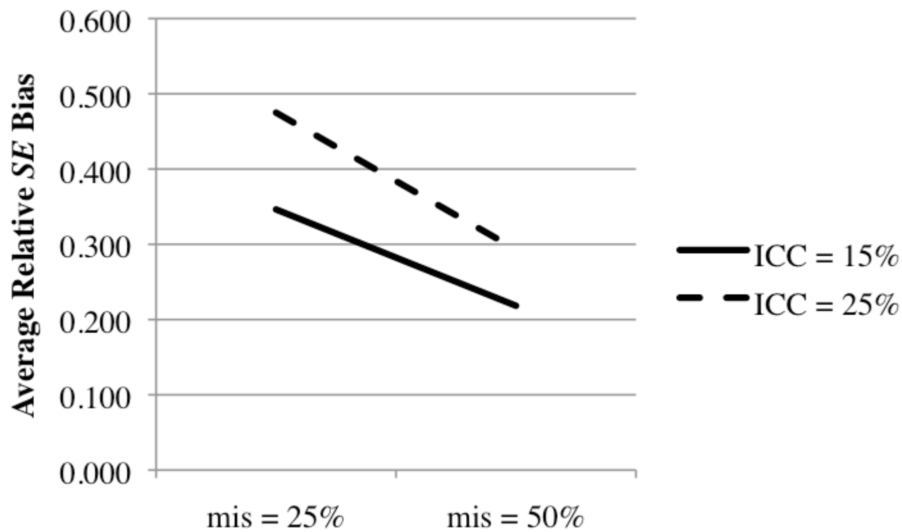

Figure 9. ICC and Percent of Mobile Students Missing an Identifier Interaction Effect on the Relative *SE* Bias of γ_{00} Estimates for MMREM-*Delete*

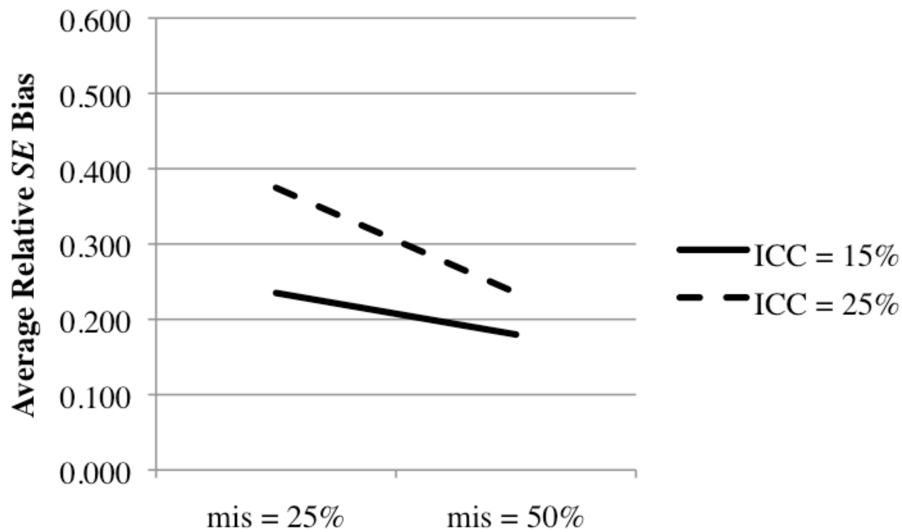

Figure 10. ICC and Percent of Mobile Students Missing an Identifier Interaction Effect on the Relative *SE* Bias of γ_{00} Estimates for HLM-*Delete*

While the ANOVA for MMREM-*Delete* only resulted in the two common two-way interactions just discussed, MMREM-*Unique* and HLM-*Delete* had additional significant two-way interactions. Both MMREM-*Unique* and HLM-*Delete* had significant mobility percentage by ICC interactions [MMREM-*Unique*: $F(1, 15989) = 1282.5$, $p < 0.001$, $\eta_p^2 = 0.074$; HLM-*Delete*: $F(1, 15989) = 455.7$, $p < 0.001$, $\eta_p^2 = 0.028$]. The trend indicated that the difference in bias was greater between ICC = 15% and ICC = 25% when the mobility percentage was 10% (MMREM-*Unique*: $M_{m=10\%,\ \mathrm{ICC}=15\%} = 0.368$, $M_{m=10\%,\ \mathrm{ICC}=25\%} = 0.510$; HLM-*Delete*: $M_{m=10\%,\ \mathrm{ICC}=15\%} = 0.253$, $M_{m=10\%,\ \mathrm{ICC}=25\%} = 0.383$) as opposed to 20% (MMREM-*Unique*: $M_{m=20\%,\ \mathrm{ICC}=15\%} = 0.332$, $M_{m=20\%,\ \mathrm{ICC}=25\%} = 0.359$; HLM-*Delete*: $M_{m=20\%,\ \mathrm{ICC}=15\%} = 0.162$, $M_{m=20\%,\ \mathrm{ICC}=25\%} = 0.226$). Depictions of this interaction are given in Figures 11 and 12. In addition to the three other significant two-way interactions, the HLM-*Delete* ANOVA results indicated a significant percentage of mobile students, *m*, by percentage of mobile students missing an identifier, *mis*, interaction [$F(1,15989) = 424.1$, $p < 0.001$, $\eta_p^2 = 0.026$]. A larger decrease in bias was found between the *mis* = 25% and *mis* = 50% conditions when *m* = 10% ($M_{m=10\%,\ mis=25\%}$ = 0.383, $M_{m=10\%,\ mis=50\%}$ = 0.254) as compared to *m* = 20% ($M_{m=20\%,\ mis=25\%}$ = 0.227, $M_{m=20\%,\ mis=50\%}$ = 0.162). Figure 13 is a depiction of this relationship.

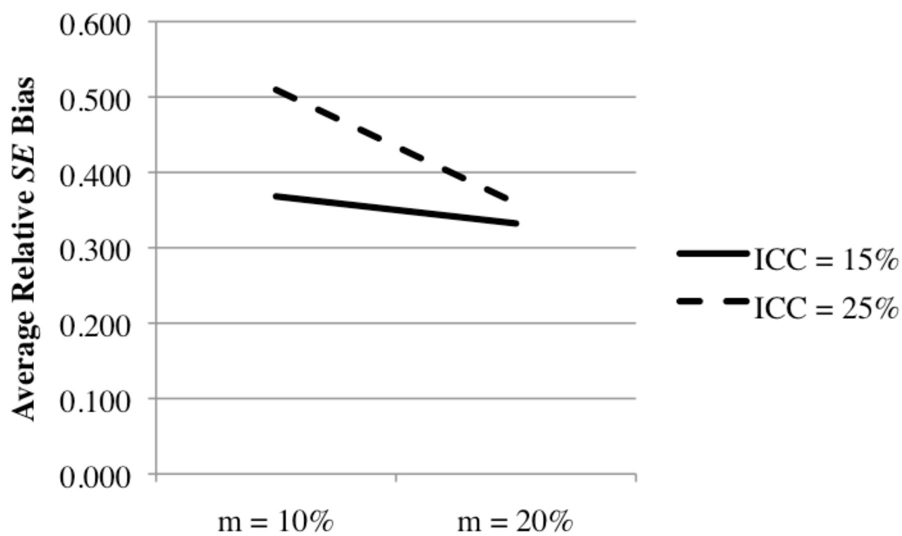

Figure 11. Mobility Percentage and ICC Interaction Effect on the Relative *SE* Bias of γ_{00} Estimates for MMREM-*Unique*

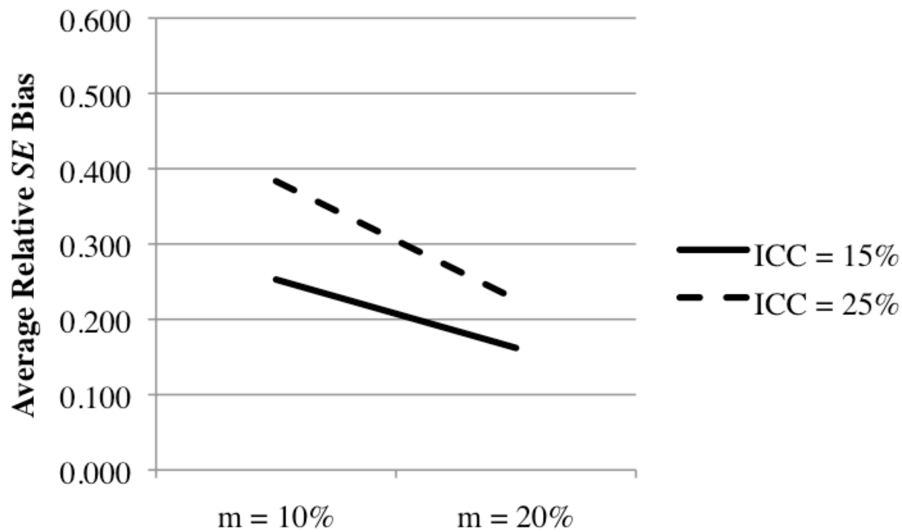

Figure 12. Mobility Percentage and ICC Interaction Effect on the Relative *SE* Bias of γ_{00} Estimates for HLM-*Delete*

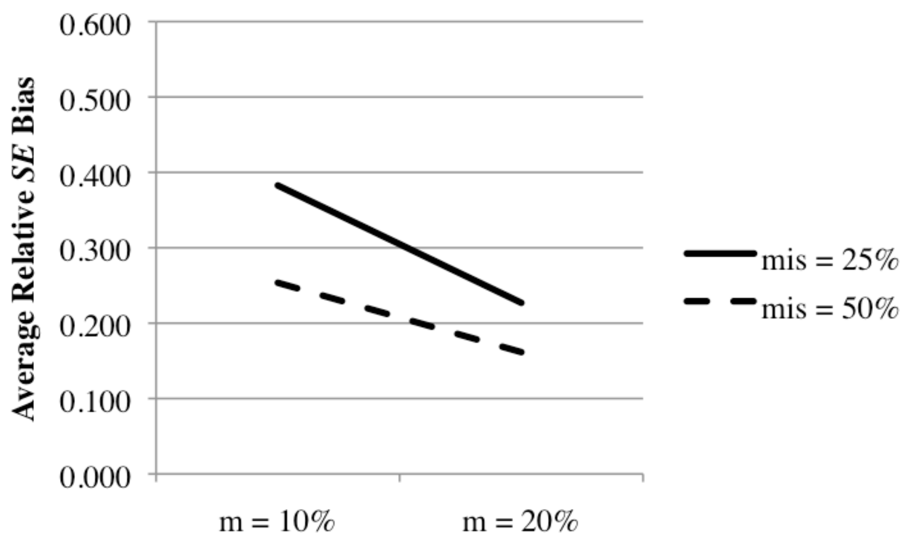

Figure 13. Mobility Percentage and Percent of Mobile Students with a Missing Identifier Interaction Effect on the Relative *SE* Bias of γ_{00} Estimates for HLM-*Delete*

X Coefficient, γ_{10}

No substantial relative *SE* bias was found for the *X* coefficient, γ_{10}, estimates (see Table 12).

Table 12

Relative Standard Error Bias of the X Coefficient, γ_{10}, Estimates

Condition				Procedure		
M	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	-0.054	-0.053	-0.042
10%	15%	50	50%	0.009	0.007	0.001
10%	15%	100	25%	0.030	0.031	0.032
10%	15%	100	50%	-0.018	-0.015	-0.017
10%	25%	50	25%	0.012	0.009	0.015
10%	25%	50	50%	0.006	0.013	0.012
10%	25%	100	25%	-0.011	-0.015	-0.029
10%	25%	100	50%	-0.011	-0.019	-0.018
20%	15%	50	25%	-0.030	-0.030	-0.017
20%	15%	50	50%	-0.026	-0.034	-0.033
20%	15%	100	25%	0.002	-0.006	-0.012
20%	15%	100	50%	0.053	0.047	0.057
20%	25%	50	25%	-0.007	0.000	-0.010
20%	25%	50	50%	-0.002	0.006	0.011
20%	25%	100	25%	0.013	0.016	0.032
20%	25%	100	50%	-0.035	-0.035	-0.026

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Mobility Coefficient, γ_{20}

Across both procedures and conditions, no substantial *SE* bias was found in the mobility coefficient, γ_{20}, estimates (see Table 13). Note that the mobility coefficient, γ_{20}, was not estimated when employing HLM-*Delete* as no mobile students were in the dataset.

Table 13

Relative Standard Error Bias of the Mobility Coefficient, γ_{20}, Estimates

Condition				Procedure	
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*
10%	15%	50	25%	0.025	0.001
10%	15%	50	50%	0.015	-0.006
10%	15%	100	25%	0.003	-0.021
10%	15%	100	50%	-0.002	0.022
10%	25%	50	25%	0.012	0.029
10%	25%	50	50%	0.032	0.000
10%	25%	100	25%	0.052	0.056
10%	25%	100	50%	0.012	0.022
20%	15%	50	25%	0.020	0.029
20%	15%	50	50%	0.017	0.025
20%	15%	100	25%	0.003	0.006
20%	15%	100	50%	-0.004	0.006
20%	25%	50	25%	0.002	0.003
20%	25%	50	50%	0.004	-0.007
20%	25%	100	25%	-0.004	-0.005
20%	25%	100	50%	0.003	0.003

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Level One Variance, σ^2

There was no substantial relative *SE* bias found for the level one variance component, σ^2, estimates (see Table 14).

Table 14

Relative Standard Error Bias of the Level One Variance Component, σ^2, Estimates

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	0.019	0.020	0.011
10%	15%	50	50%	0.021	0.027	0.021
10%	15%	100	25%	-0.004	0.003	0.005
10%	15%	100	50%	0.027	0.027	0.013
10%	25%	50	25%	0.027	0.037	0.043
10%	25%	50	50%	0.016	0.024	0.015
10%	25%	100	25%	0.014	0.019	-0.004
10%	25%	100	50%	-0.021	-0.014	-0.018
20%	15%	50	25%	0.047	0.058	0.046
20%	15%	50	50%	-0.006	0.011	0.016
20%	15%	100	25%	-0.002	-0.007	-0.003
20%	15%	100	50%	-0.003	0.004	0.000
20%	25%	50	25%	-0.007	0.007	0.028
20%	25%	50	50%	-0.001	-0.013	-0.033
20%	25%	100	25%	-0.029	-0.031	-0.025
20%	25%	100	50%	-0.003	0.016	0.019

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Level Two Variance, τ_{00}

The relative *SE* bias for the estimates of the level two variance component, τ_{00}, is shown in Table 15. Substantial positive *SE* bias was found across all conditions for estimates resulting from use of the MMREM-*Unique* procedure, 14 out of the 16 conditions with the MMREM-*Delete* procedure, and 12 out of the 16 conditions for the HLM-*Delete* procedure. HLM-*Delete* estimates consistently had the least amount of bias when compared with the other procedures with the exception of one condition (m = 10%, ICC = 25%, c = 50, *mis* = 50%) where MMREM-*Delete* and HLM-*Delete* estimates resulted in the same amount of bias to three decimal places. The bias for MMREM-*Unique* ranged from 12.9% to 44.9% with an average of 24.5%, MMREM-*Delete* ranged from 8.3% to 42.3% with a mean of 21.5%, and HLM-*Delete* ranged from 5.2% to 35.4% with an average of 16.9%.

Table 15

Relative Standard Error Bias of the Level Two Variance Component, τ_{00}, Estimates

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	***0.272***	***0.250***	***0.158***
10%	15%	50	50%	***0.236***	***0.207***	***0.173***
10%	15%	100	25%	***0.224***	***0.207***	***0.155***
10%	15%	100	50%	***0.200***	***0.173***	***0.139***
10%	25%	50	25%	***0.449***	***0.423***	***0.354***
10%	25%	50	50%	***0.317***	***0.279***	***0.279***
10%	25%	100	25%	***0.358***	***0.342***	***0.299***
10%	25%	100	50%	***0.251***	***0.217***	***0.189***
20%	15%	50	25%	***0.239***	***0.193***	0.100
20%	15%	50	50%	***0.149***	0.097	0.061
20%	15%	100	25%	***0.171***	***0.153***	***0.106***
20%	15%	100	50%	***0.129***	0.083	0.052
20%	25%	50	25%	***0.300***	***0.278***	***0.247***
20%	25%	50	50%	***0.219***	***0.185***	***0.163***
20%	25%	100	25%	***0.266***	***0.251***	***0.169***
20%	25%	100	50%	***0.135***	***0.103***	0.055

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

All four manipulated conditions significantly affected the *SE* bias in all procedures. In the case of the mobility percentage [MMREM-*Unique*: $F(1, 15989) = 928.1$, $p < 0.001$, $\eta_p^2 = 0.055$; MMREM-*Delete*: $F(1, 15989) = 1081.7$, $p < 0.001$, $\eta_p^2 =$

0.063; HLM-*Delete*: $F(1, 15989) = 1192.3$, $p < 0.001$, $\eta_p^2 = 0.069$], as the percent of mobile students, *m*, increased, the bias decreased (MMREM-*Unique*: $M_{m=10\%} = 0.288$, $M_{m=20\%} = 0.201$; MMREM-*Delete*: $M_{m=10\%} = 0.262$, $M_{m=20\%} = 0.168$; HLM-*Delete*: $M_{m=10\%} = 0.218$, $M_{m=20\%} = 0.119$). Similarly, the main effect of the percent of mobile students that had a missing identifier was significantly related to the relative *SE* bias [MMREM-*Unique*: $F(1, 15989) = 783.5$, $p < 0.001$, $\eta_p^2 = 0.047$; MMREM-*Delete*: $F(1, 15989) = 1088.0$, $p < 0.001$, $\eta_p^2 = 0.064$; HLM-*Delete*: $F(1, 15989) = 431.7$, $p < 0.001$, $\eta_p^2 = 0.026$]. Similar to the mobility trend, as the percent of mobile students that had a missing school identifier, *mis*, increased, the bias decreased (MMREM-*Unique*: $M_{mis=25\%} = 0.285$, $M_{mis=50\%} = 0.205$; MMREM-*Delete*: $M_{mis=25\%} = 0.262$, $M_{mis=50\%} = 0.168$; HLM-*Delete*: $M_{mis=25\%} = 0.198$, $M_{mis=50\%} = 0.139$). The ICC also significantly impacted the relative *SE* bias across all procedures [MMREM-*Unique*: $F(1, 15989) = 866.7$, $p < 0.001$, $\eta_p^2 = 0.051$; MMREM-*Delete*: $F(1, 15989) = 983.7$, $p < 0.001$, $\eta_p^2 = 0.058$; HLM-*Delete*: $F(1, 15989) = 1254.8$, $p < 0.001$, $\eta_p^2 = 0.073$]. The relative *SE* bias increased with an increase in the ICC value (MMREM-*Unique*: $M_{\text{ICC}=15\%} = 0.203$, $M_{\text{ICC}=25\%} = 0.287$; MMREM-*Delete*: $M_{\text{ICC}=15\%} = 0.170$, $M_{\text{ICC}=25\%} = 0.260$; HLM-*Delete*: $M_{\text{ICC}=15\%} = 0.118$, $M_{\text{ICC}=25\%} = 0.219$). The final manipulated condition, number of level two units, also affected the relative *SE* bias significantly [MMREM-*Unique*: $F(1, 15989) = 378.9$, $p < 0.001$, $\eta_p^2 = 0.023$; MMREM-*Delete*: $F(1, 15989) = 279.7$, $p < 0.001$, $\eta_p^2 = 0.017$; HLM-*Delete*: $F(1, 15989) = 259.4$, $p < 0.001$, $\eta_p^2 = 0.016$] such that when the number of level two units, *c*, increased from 50 to 100, the relative *SE* bias decreased (MMREM-*Unique*: $M_{c=50}$ = 0.273, $M_{c=100}$ = 0.217; MMREM-*Delete*: $M_{c=50}$ = 0.239, $M_{c=100}$ = 0.191; HLM-*Delete*: $M_{c=50}$ = 0.192, $M_{c=100}$ = 0.146). No two-way interactions were significant for any procedure investigated.

Chapter 5: Discussion

This study compared three different ad hoc procedures for handling missing level two identifiers when a multiple membership data structure is present. MMREM-*Unique* replaced the missing identifier with a pseudo-identifier while MMREM-*Delete* removed cases with a missing identifier and HLM-*Delete* removed any mobile subject.

Substantial relative parameter bias was found for estimates of all parameters with the exception of the intercept, γ_{00}. Substantial relative standard error bias was found for estimates of the intercept, γ_{00}, and for the level two variance component, τ_{00}. Details of these findings are discussed in this chapter.

Relative Parameter Bias

Substantial relative parameter bias was found for the estimates of the coefficient of X, γ_{10} (see Table 7); the coefficient of M, γ_{20} (see Table 8); the level one variance component, σ^2 (see Table 9); and the level two variance component, τ_{00} (see Table 10). No substantial relative parameter bias was found for estimates of the intercept, γ_{00}, (see Table 6) thus no further discussion will be presented on the intercept in regards to the relative parameter bias.

X Coefficient, γ_{10}

All three procedures performed similarly in regards to the estimates of the X coefficient, γ_{10} (see Table 7). The relative parameter bias values did not differ more than 0.4% across procedures and ranged from an overestimation of 36.3% to 38.4%. Previous research (Chung & Beretvas, 2012; Wolff Smith & Beretvas, 2011) did not find substantial relative parameter bias for a similar predictor when employing either MMREM or HLM. To investigate whether the overestimation found was due to the misspecification of the model through the omission of the variable P, the true generating model (Equation 24) was also estimated using the generated datasets. Remember that when generating the data, all school identifiers were present. Those students who were deemed mobile and missing an identifier were assigned a level two unit outside of the 50 or 100 schools (condition-dependent) indicating that the school was outside of the study's focal area. From here on, this procedure will be termed, MMREM-*True*. The relative parameter bias of the X coefficient estimates for the four procedures is shown in Table 16.

Table 16

Relative Parameter Bias of the X Coefficient, γ_{10}, Estimates for Four Procedures

Condition				Procedure			
m	ICC	*c*	*mis*	MMREM *True*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	-0.003	***0.376***	***0.376***	***0.376***
10%	15%	50	50%	0.001	***0.381***	***0.381***	***0.381***
10%	15%	100	25%	0.001	***0.382***	***0.382***	***0.384***
10%	15%	100	50%	0.000	***0.381***	***0.381***	***0.382***
10%	25%	50	25%	0.002	***0.382***	***0.381***	***0.381***
10%	25%	50	50%	-0.002	***0.377***	***0.377***	***0.378***
10%	25%	100	25%	0.000	***0.380***	***0.380***	***0.381***
10%	25%	100	50%	0.000	***0.380***	***0.380***	***0.382***
20%	15%	50	25%	-0.001	***0.365***	***0.364***	***0.368***
20%	15%	50	50%	0.002	***0.368***	***0.366***	***0.368***
20%	15%	100	25%	-0.001	***0.366***	***0.363***	***0.366***
20%	15%	100	50%	0.001	***0.366***	***0.364***	***0.367***
20%	25%	50	25%	0.000	***0.367***	***0.365***	***0.369***
20%	25%	50	50%	0.000	***0.367***	***0.365***	***0.369***
20%	25%	100	25%	0.003	***0.367***	***0.365***	***0.369***
20%	25%	100	50%	0.002	***0.367***	***0.364***	***0.368***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Given the fact that no bias was found when estimating MMREM-*True*, the overestimation appears to be a result of the model misspecification. The level one predictor, *X*, was generated to be a function of the variable *P* that was not included in any of the models estimated (except for MMREM-*True*). Omission of *P* from the estimating models resulted in some of the explanatory power of *P* being attributed to the modeled *X* variable.

With the cutoff used ($\eta_p^2 \geq 0.01$), the ANOVA results indicated that the mobility percentage (*m*) significantly affected the severity of the bias in the estimation of the *X* coefficient, γ_{10}, in each procedure. However, upon inspection of the relative parameter bias means, the largest difference in the average bias between the $m = 10\%$ and $m = 20\%$ conditions was 1.6% which is less than a 5% difference. Even though the η_p^2 value indicated this is a practically significant difference, the effect is small.

M Coefficient, γ_{20}

Very positive substantial bias, on the order of 700%, was found in both the MMREM-*Unique* and MMREM-*Delete* estimates of the mobility (*M*) coefficient, γ_{20} (see Table 8) indicating that the estimates were of substantially larger magnitude than the true value. In this case, the estimates of mobility's effect were more negative than the true effect (which was negative). This parameter was not estimated when HLM-*Delete* was employed as the procedure removes mobile students from the analytic dataset. Across all conditions, MMREM-*Delete* resulted in less positive relative parameter bias as compared with MMREM-*Unique* ($M_{\text{MMREM-}Delete} = 6.865$, $M_{\text{MMREM-}Unique} = 7.984$). Under the MMREM-*Delete* procedure for handling missing identifiers, data for mobile students with missing identifiers was removed from the dataset. As such, the effect of mobility is not unexpectedly smaller because the students who have been generated to be the worst performing (as was done in the current study because mobile students with missing identifiers were generated to have the lowest outcome scores) have been removed. Wolff Smith and Beretvas (2011) similarly found substantial positive relative parameter bias for the coefficient of a level one mobility predictor when estimating both an MMREM and HLM where the MMREM matched the generating model. Neither of these models, however, resulted in bias as extreme as was found in this study. In the previous study, however, mobility was generated randomly. Thus, the mobile students were not necessarily the lower performing students. About half of the conditions (even under conditions with the correct model being estimated and no missing identifiers) resulted in some substantial bias indicating that the effect of mobility is hard to estimate even when no omitted variable bias or missing level two identifiers are present.

To investigate whether or not this substantial bias is a result of model misspecification, the relative parameter bias in the γ_{20} estimates was also assessed using the MMREM-*True* procedure. The relative parameter bias of the estimates of γ_{20} can be found in Table 17. Substantial positive relative parameter bias was found for estimates of the *M* coefficient when MMREM-*True* was employed matching what would be expected given Wolff Smith and Beretvas' (2011) findings. This shows that the level one mobility effect is difficult to estimate even when no misspecification is present. However, the average bias was not nearly as substantial using MMREM-*True* ($M_{\text{MMREM-}True}$ = 0.677) as under the other procedures (MMREM-*Delete* and MMREM-*Unique*). The ratio of the degree of bias between MMREM-*True* estimates and those of the other two procedures (MMREM-*Unique* and MMREM-*Delete*) was approximately one to seven.

Table 17

Relative Parameter Bias of the M Coefficient, γ_{20}, Estimates for Three Procedures

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *True*	MMREM *Unique*	MMREM *Delete*
10%	15%	50	25%	***0.282***	***7.586***	***7.011***
10%	15%	50	50%	***0.471***	***7.877***	***6.935***
10%	15%	100	25%	***0.515***	***7.799***	***7.026***
10%	15%	100	50%	***1.064***	***8.453***	***7.137***
10%	25%	50	25%	***0.689***	***7.954***	***7.267***
10%	25%	50	50%	***0.979***	***8.370***	***7.080***
10%	25%	100	25%	***0.822***	***8.230***	***7.236***
10%	25%	100	50%	***1.413***	***8.697***	***6.675***
20%	15%	50	25%	***0.235***	***7.498***	***6.857***
20%	15%	50	50%	***0.506***	***7.804***	***6.563***
20%	15%	100	25%	***0.368***	***7.658***	***6.823***
20%	15%	100	50%	***0.872***	***8.106***	***6.546***
20%	25%	50	25%	***0.225***	***7.565***	***6.896***
20%	25%	50	50%	***0.571***	***7.813***	***6.545***
20%	25%	100	25%	***0.590***	***7.900***	***6.887***
20%	25%	100	50%	***1.226***	***8.426***	***6.361***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

The true intercept value was set at 100 with mobility modeled as having a small effect (γ_{20} = -0.5). Wolff Smith and Beretvas (2011) considered the influence of the coefficient value. In their study, the value of the level two mobility predictor was manipulated. They found that the value did have an effect on some of the fixed effects estimates. As such, it seemed possible that the true value of the *M* coefficient might have affected parameter recovery such that the larger effect, the less the bias. To test this hypothesis, additional conditions were generated in which the true value of the *M* coefficient was set to −5 (instead of −0.5). The relative parameter bias results of this modification are displayed in Table 18. With conditions where the coefficient is larger in magnitude (meaning the effect of mobility is stronger), the parameter is better estimated. Five of the sixteen conditions for MMREM-*True* are now unbiased. Additionally, the relative parameter bias has decreased by a factor of approximately ten across procedures, the same factor of which the value of the coefficient has increased. It is noteworthy to point out that the ratio of the amount of relative parameter bias for MMREM-*True* estimates versus those using the other two procedures is still approximately one to seven. Given this ad hoc analysis, it appears that the bias is directly a function of the mobility coefficient's true value such that small effects are harder to estimate and thus result in more biased estimates.

Table 18

Relative Parameter Bias of the M Coefficient, γ_{20}, Estimates (γ_{20} = -5)

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *True*	MMREM *Unique*	MMREM *Delete*
10%	15%	50	25%	0.028	***0.759***	***0.702***
10%	15%	50	50%	0.047	***0.788***	***0.695***
10%	15%	100	25%	***0.052***	***0.780***	***0.703***
10%	15%	100	50%	***0.106***	***0.845***	***0.715***
10%	25%	50	25%	***0.069***	***0.795***	***0.728***
10%	25%	50	50%	***0.098***	***0.837***	***0.710***
10%	25%	100	25%	***0.082***	***0.823***	***0.724***
10%	25%	100	50%	***0.141***	***0.870***	***0.669***
20%	15%	50	25%	0.023	***0.750***	***0.687***
20%	15%	50	50%	***0.051***	***0.780***	***0.659***
20%	15%	100	25%	0.037	***0.766***	***0.683***
20%	15%	100	50%	***0.087***	***0.811***	***0.656***
20%	25%	50	25%	0.023	***0.757***	***0.691***
20%	25%	50	50%	***0.057***	***0.781***	***0.658***
20%	25%	100	25%	***0.059***	***0.790***	***0.690***
20%	25%	100	50%	***0.123***	***0.843***	***0.638***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Level One Variance, σ^2

The level one variance component, σ^2, estimates were substantially positively biased across procedures. All three procedures (MMREM-*Unique*, MMREM-*Delete*, and

HLM-*Delete*) were very similar in their bias only differing by, at most, 0.6%. No condition was found to substantially impact the relative parameter bias across procedures. Chung and Beretvas (2012) as well as Meyers and Beretvas (2006) similarly found that the level one variance component was overestimated when HLM was employed. However, contrary to the results here, no substantial parameter bias was found in cases where the MMREM (Chung & Beretvas, 2012) or CCREM (Meyers & Beretvas, 2006) was used.

To determine if the level one variance component was not estimated well or if this overestimation was due to the model misspecification, MMREM-*True* was estimated and the relative parameter bias for the level one variance estimates was calculated (see Table 19). No substantial relative parameter bias was found in any condition when MMREM-*True* was estimated. This is evidence that the overestimation of this parameter is due to the omission of *P* from the models being estimated. The unique relationship between *P* and *Y* is unexplained as *P* was not included in the models estimated using the MMREM-*Unique,* MMREM-*Delete*, or HLM-*Delete* procedures. As such, the level one variance has increased to account for the resulting unexplained level one variability.

Table 19

Relative Parameter Bias of the Level One Variance Component, σ^2, Estimates for Four Procedures

Condition				Procedure			
m	ICC	*c*	*mis*	MMREM *True*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	0.004	***0.420***	***0.420***	***0.423***
10%	15%	50	50%	0.002	***0.416***	***0.416***	***0.418***
10%	15%	100	25%	0.001	***0.415***	***0.415***	***0.416***
10%	15%	100	50%	0.000	***0.414***	***0.414***	***0.415***
10%	25%	50	25%	0.002	***0.414***	***0.414***	***0.416***
10%	25%	50	50%	0.001	***0.416***	***0.416***	***0.417***
10%	25%	100	25%	-0.001	***0.412***	***0.412***	***0.413***
10%	25%	100	50%	0.000	***0.415***	***0.415***	***0.415***
20%	15%	50	25%	0.005	***0.414***	***0.414***	***0.420***
20%	15%	50	50%	0.003	***0.415***	***0.417***	***0.420***
20%	15%	100	25%	0.000	***0.410***	***0.409***	***0.411***
20%	15%	100	50%	0.002	***0.411***	***0.412***	***0.414***
20%	25%	50	25%	0.002	***0.413***	***0.413***	***0.417***
20%	25%	50	50%	0.000	***0.412***	***0.412***	***0.416***
20%	25%	100	25%	0.002	***0.410***	***0.409***	***0.411***
20%	25%	100	50%	0.004	***0.414***	***0.412***	***0.415***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

Level Two Variance, τ_{00}

The level two variance component, τ_{00}, was found to be substantially underestimated across all conditions when MMREM-*Delete* and HLM-*Delete* were employed while substantial underestimation was found for 12 of the 16 conditions when MMREM-*Unique* was employed. Across all conditions, estimates using MMREM-*Unique* resulted in the least amount of bias. For the most part, these results are contrary to what has been found in previous research. Both Chung and Beretvas (2012) and Wolff Smith and Beretvas (2011) found this variance component to be overestimated when estimating the correct MMREM. Additionally, when employing HLM after deleting all mobile students' data from the analysis, Wolff Smith and Beretvas (2011) found substantial positive bias in the estimation of τ_{00}. Neither of these studies included the added complication of model misspecification due to the omission of a variable. Additionally, no level two identifiers were missing and mobility was generated randomly. Thus, it is thought that the added misspecification (removal of *P*), missing level two identifiers, and random mobility generation might lie at the source of some of this discrepancy.

To explore whether the model misspecification contributes to the substantial relative parameter bias of the level two variance component estimates, MMREM-*True* was estimated across all conditions and the resulting relative parameter bias is shown in Table 20. Only two conditions (m = 20%, ICC = 15%, c = 50, *mis* = 50%; m = 20%, ICC = 25%, c = 50, *mis* = 50%) resulted in a slightly substantial overestimation (0.060 and 0.071, respectively) of the variance component. This gives evidence that the model misspecification might lie at the root of the substantial relative parameter bias of τ_{00}. This means that if a related variable is omitted from the model, it is likely that the level two variance component might be substantially underestimated. If, however, the related variable is included, the level two variance component will be more accurately estimated. In practice, however, it is difficult to know what variables to include or exclude from the model and how those decisions will affect the bias. In the field of econometrics, Clarke (2005) has justly called omitted variable bias a “phantom menace” and suggests that models created are first-order approximations at best.

Table 20

Relative Parameter Bias of the Level Two Variance Component, τ_{00}, Estimates for Four Procedures

Condition				Procedure			
m	ICC	*c*	*mis*	MMREM *True*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	0.019	***-0.113***	***-0.135***	***-0.181***
10%	15%	50	50%	0.019	***-0.113***	***-0.157***	***-0.186***
10%	15%	100	25%	-0.011	***-0.129***	***-0.150***	***-0.199***
10%	15%	100	50%	-0.028	***-0.148***	***-0.188***	***-0.219***
10%	25%	50	25%	0.022	-0.045	***-0.067***	***-0.113***
10%	25%	50	50%	0.032	-0.036	***-0.076***	***-0.104***
10%	25%	100	25%	-0.026	***-0.091***	***-0.109***	***-0.148***
10%	25%	100	50%	-0.043	***-0.109***	***-0.142***	***-0.166***
20%	15%	50	25%	0.047	***-0.087***	***-0.120***	***-0.184***
20%	15%	50	50%	***0.060***	***-0.079***	***-0.142***	***-0.171***
20%	15%	100	25%	0.000	***-0.122***	***-0.151***	***-0.213***
20%	15%	100	50%	-0.015	***-0.141***	***-0.195***	***-0.231***
20%	25%	50	25%	0.045	-0.023	***-0.054***	***-0.123***
20%	25%	50	50%	***0.071***	0.003	***-0.057***	***-0.101***
20%	25%	100	25%	-0.029	***-0.093***	***-0.118***	***-0.171***
20%	25%	100	50%	-0.028	***-0.096***	***-0.145***	***-0.182***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

The ANOVA results indicated that across all procedures, the true ICC value and number of level two units, *c*, were significantly related to the relative parameter bias. As the true ICC value increased, the bias decreased (bias became less negative). In other words, as the proportion of variance attributed to the level two units increased, the bias decreased. No previous research has found a substantial impact of the true ICC value on the relative parameter bias of the level two variance component, τ_{00}. Future research should explore this relationship further.

The trend that an increase in level two units (c = 50 to c = 100) results in more bias as found in this study was also found by Chung and Beretvas (2012), but can seem counterintuitive. However, it seems possible that two counteracting mechanisms might be at play. Browne and Draper (2006) found that for datasets where fewer level two units (smaller values of c) are present, variance components at the highest level (here, level two) are overestimated when MCMC estimation is employed. While the Browne and Draper (2006) conclusions indicate that substantial positive bias should be found in the level two variance component estimates when a smaller number of level two units is present, substantial negative bias was found here. Given the resulting overall substantial negative relative parameter bias for the level two variance component in this study, it appears that the model misspecification results in some negative bias. The negative bias decreases when the number of level two units increases (c = 50 to c = 100), but the positive bias decreases at a faster rate. As such, the overall bias will appear to worsen with an increase in the number of level two units.

To further investigate the counteracting mechanism hypothesis regarding the level two units, the full set of conditions was run with 200 level two units. The results of this analysis are shown in Table 21. It was found that the bias decreased with an increase in level two units from c = 100 to c = 200 ($M_{c=50}$ = -0.103, $M_{c=100}$ = -0.152, $M_{c=200}$ = -0.124) giving evidence that this hypothesis might be feasible. It is thought that the positive bias due to the small number of level two units is about the same once c = 100 and the negative bias due to model misspecification decreases as c increases. Between c = 100 and c = 200, the only bias changing is the negative bias due to the model misspecification thus resulting in less negative bias when c = 200 as compared to when c = 100. Therefore, the counteracting mechanism hypothesis appears to be reasonable.

Table 21

Relative Parameter Bias of the Level Two Variance Component, τ_{00}, Estimates With c = 200

Condition				Procedure		
m	ICC	*c*	*mis*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	***-0.113***	***-0.135***	***-0.181***
10%	15%	50	50%	***-0.113***	***-0.157***	***-0.186***
10%	15%	100	25%	***-0.129***	***-0.150***	***-0.199***
10%	15%	100	50%	***-0.148***	***-0.188***	***-0.219***
10%	15%	200	25%	***-0.128***	***-0.148***	***-0.197***
10%	15%	200	50%	***-0.125***	***-0.163***	***-0.197***
10%	25%	50	25%	-0.045	***-0.067***	***-0.113***
10%	25%	50	50%	-0.036	***-0.076***	***-0.104***
10%	25%	100	25%	***-0.091***	***-0.109***	***-0.148***
10%	25%	100	50%	***-0.109***	***-0.142***	***-0.166***
10%	25%	200	25%	***-0.068***	***-0.088***	***-0.136***
10%	25%	200	50%	***-0.069***	***-0.106***	***-0.136***
20%	15%	50	25%	***-0.087***	***-0.120***	***-0.184***
20%	15%	50	50%	***-0.079***	***-0.142***	***-0.171***
20%	15%	100	25%	***-0.122***	***-0.151***	***-0.213***
20%	15%	100	50%	***-0.141***	***-0.195***	***-0.231***
20%	15%	200	25%	***-0.110***	***-0.136***	***-0.191***
20%	15%	200	50%	***-0.098***	***-0.147***	***-0.179***
20%	25%	50	25%	-0.023	***-0.054***	***-0.123***
20%	25%	50	50%	0.003	***-0.057***	***-0.101***
20%	25%	100	25%	***-0.093***	***-0.118***	***-0.171***
20%	25%	100	50%	***-0.096***	***-0.145***	***-0.182***
20%	25%	200	25%	***-0.052***	***-0.079***	***-0.145***
20%	25%	200	50%	-0.041	***-0.095***	***-0.137***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

RELATIVE STANDARD ERROR BIAS

Substantial relative standard error (*SE*) bias was found for estimates of the intercept, γ_{00} (see Table 11), and level two variance component, τ_{00} (see Table 15).

Standard error estimates of the *X* coefficient, γ_{10} (see Table 12); *M* coefficient, γ_{20} (see Table 13); and level one variance component, σ^2 (see Table 14); were not substantially biased and thus will not be discussed further.

Intercept, γ_{00}

Substantial positive relative *SE* bias was found across all conditions and procedures (MMREM-*Unique*, MMREM-*Delete*, HLM-*Delete*) indicating an overestimation of the *SE* of the intercept. HLM-*Delete* resulted in the smallest amount of bias with MMREM-*Unique* resulting in the most relative *SE* bias. Previous research with the MMREM has not found *SE* bias in the estimates of the intercept (Wolff Smith & Beretvas, 2011), although some research on estimates of the CCREM did find that the intercept's *SE* was underestimated (Luo & Kwok, 2012).

To investigate whether or not this bias was a result of model misspecification, MMREM-*True* was estimated and the relative *SE* bias of the intercept estimates was calculated (see Table 22). All conditions resulted in an overestimation of the intercept estimate's *SE* with the bias ranging from 25.2% to 68.6%. Thus, it does not appear that the model misspecification is at the root of this overestimation.

Table 22

Relative SE Bias of the Intercept, γ_{00}, Estimates for Four Procedures

Condition				Procedure			
m	ICC	*c*	*mis*	MMREM *True*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	***0.396***	***0.391***	***0.377***	***0.282***
10%	15%	50	50%	***0.308***	***0.327***	***0.241***	***0.204***
10%	15%	100	25%	***0.501***	***0.460***	***0.445***	***0.317***
10%	15%	100	50%	***0.275***	***0.296***	***0.277***	***0.208***
10%	25%	50	25%	***0.686***	***0.663***	***0.534***	***0.461***
10%	25%	50	50%	***0.429***	***0.435***	***0.377***	***0.328***
10%	25%	100	25%	***0.600***	***0.562***	***0.573***	***0.470***
10%	25%	100	50%	***0.425***	***0.378***	***0.341***	***0.274***
20%	15%	50	25%	***0.352***	***0.373***	***0.260***	***0.149***
20%	15%	50	50%	***0.276***	***0.294***	***0.148***	***0.113***
20%	15%	100	25%	***0.413***	***0.379***	***0.303***	***0.192***
20%	15%	100	50%	***0.275***	***0.283***	***0.207***	***0.195***
20%	25%	50	25%	***0.515***	***0.461***	***0.413***	***0.266***
20%	25%	50	50%	***0.296***	***0.307***	***0.250***	***0.193***
20%	25%	100	25%	***0.438***	***0.429***	***0.379***	***0.301***
20%	25%	100	50%	***0.252***	***0.240***	***0.204***	***0.146***

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

The ANOVAs were run to determine which condition(s) were significantly related to the relative *SE* bias. Various main effects and two-way interactions were found to be practically significant ($\eta_p^2 \geq 0.01$). Three conditions' main effects (mobility percentage, ICC, and percent of mobile students missing an identifier) were practically significant across all three procedures (MMREM-*Unique*, MMREM-*Delete*, HLM-*Delete*). Additionally, the main effect of the number of level two units was practically significant for MMREM-*Unique*. Even though the main effect of the number of level two units met the threshold set for practical significance ($\eta_p^2 \geq 0.01$), the difference in mean relative *SE* relative bias was less than 3%. Luo and Kwok (2012) similarly found that the mobility rate and number of level two units had a large effect on the relative *SE* bias of the intercept. These condition main effects were dependent on another condition as evidenced by practically significant two-way interactions.

While many two-way interactions were found to be practically significant ($\eta_p^2 \geq 0.01$), none of them had a large effect using Cohen's (1977) criteria ($\eta_p^2 > 0.14$). The majority of the interactions had small effects ($0.017 \leq \eta_p^2 \leq 0.045$) while one had a moderate effect ($\eta_p^2 = 0.074$). These interactions were found to be practically significant here using the conservative "small" effect size cutoff value, though no previous research has reported substantial two-way interactions for the relative *SE* bias of the intercept. Previous research, however, has not explored cases where missing level two identifiers are present and mobility is generated randomly when using a two-level model. As such, these additional complications may be contributing to the results found here. At present, it is unclear why these two-way interactions were found to be practically significant in this study. Future research could explore this further with additional condition manipulations.

Level Two Variance, τ_{00}

Substantial positive relative *SE* bias was found across all three procedures with the exception of two conditions when using MMREM-*Delete* and four conditions when using HLM-*Delete*. Across all conditions, HLM-*Delete* resulted in the least amount of bias while MMREM-*Unique* resulted in the most bias. While substantial bias was found across all conditions here, previous research has found either no (Luo & Kwok, 2012) or very minimal (Wolff Smith & Beretvas, 2011) relative *SE* bias for the level two variance component.

To explore whether this substantial relative *SE* bias was a result of model misspecification, the level two variance component was estimated and the relative *SE* bias was calculated when MMREM-*True* was employed (see Table 23). The results in Table 23 show that use of MMREM-*True* resulted in very slightly more positive bias than any of the other three procedures indicating that the substantial bias was not a result of model misspecification.

Table 23

Relative SE Bias of the Level Two Variance Component, τ_{00}, Estimates for Four Procedures

Condition				Procedure			
m	ICC	*c*	*mis*	MMREM *True*	MMREM *Unique*	MMREM *Delete*	HLM *Delete*
10%	15%	50	25%	***0.298***	***0.272***	***0.250***	***0.158***
10%	15%	50	50%	***0.237***	***0.236***	***0.207***	***0.173***
10%	15%	100	25%	***0.235***	***0.224***	***0.207***	***0.155***
10%	15%	100	50%	***0.206***	***0.200***	***0.173***	***0.139***
10%	25%	50	25%	***0.453***	***0.449***	***0.423***	***0.354***
10%	25%	50	50%	***0.327***	***0.317***	***0.279***	***0.279***
10%	25%	100	25%	***0.356***	***0.358***	***0.342***	***0.299***
10%	25%	100	50%	***0.254***	***0.251***	***0.217***	***0.189***
20%	15%	50	25%	***0.273***	***0.239***	***0.193***	0.100
20%	15%	50	50%	***0.166***	***0.149***	0.097	0.061
20%	15%	100	25%	***0.177***	***0.171***	***0.153***	***0.106***
20%	15%	100	50%	***0.140***	***0.129***	0.083	0.052
20%	25%	50	25%	***0.309***	***0.300***	***0.278***	***0.247***
20%	25%	50	50%	***0.227***	***0.219***	***0.185***	***0.163***
20%	25%	100	25%	***0.273***	***0.266***	***0.251***	***0.169***
20%	25%	100	50%	***0.144***	***0.135***	***0.103***	0.055

Note. *m* = percent mobility; *c* = number of level two units; *mis* = percent of mobile students missing a level two identifier; highlighted values indicate substantial bias.

The main effects of all four conditions were found to be practically significant in relation to the relative *SE* bias while no two-way interactions were significant. As expected, as the number of level two units increased from 50 to 100, the *SE* bias decreased. In all cases, as the percent of mobile students increased, the average bias decreased by between 8.7% and 9.9%. This is a counterintuitive result. One would reason that mishandling the effect of mobility in datasets with a higher proportion of mobile students would likely cause more estimation problems rather than fewer. Additionally, the bias became more positive as the ICC increased from 15% to 25%. Finally, as the percent of mobile students missing an identifier increased from 25% to 50%, the bias decreased. This seems to be a counterintuitive result for the same reason that the trend in the mobility percentage was thought to be counterintuitive. The reason for the discrepancies in the intuitive trend and actual trends is unknown and, as such, could be explored in future research.

Summary of Standard Error Bias Findings

When substantial relative *SE* bias was found, the *SE*s were overestimated. Using frequentist estimation techniques, conservative hypothesis tests would ensue from this substantial positive bias when considering performance of the *SE*s' estimation in isolation. In the case of the intercept, γ_{00}, no relative parameter bias was found. Thus, the positive *SE* bias would decrease the associated test statistics resulting in conservative hypothesis test results. The reality is, however, that applied researchers tend to not be interested in interpreting or testing the intercept parameter. Instead, applied researchers are more interested in the values and tests of the effects of each independent variable as well as the variance component parameters. In the case of the level two variance component, τ_{00}, the parameter was underestimated by an average of 12.8% across procedures while the *SE* was overestimated by an average of 20.9% across procedures. This should result in test statistics being typically underestimated and thus associated with conservative Type I error rates. When using Bayesian estimation techniques, however, credible intervals or high density regions are used (see, for example, Kruschke, 2011). These intervals or regions do not use the estimated *SE* values when testing hypotheses about parameters' values. Thus, when using Bayesian techniques for analysis, the fact that the *SE*s are biased might be of less importance for hypothesis testing.

RESULT OF GENERATING ENDOGENEITY IN MOBILITY

Wolff Smith and Beretvas (2011) found the HLM resulting from a reduced dataset in which only non-mobile students were analyzed and MMREM performed similarly. In their discussion, they hypothesized that this was the case because mobility had been randomly assigned. As such, when removing students, a random sample was removed from the dataset. In this study, mobility was not randomly generated, rather was generated as a function of omitted student level characteristics. Results of this study indicated that HLM-*Delete* and the MMREM procedures performed differently. It is then concluded that under the seemingly more reasonable scenario under which mobility is endogenous leads to differences in the performance of these models.

LIMITATIONS AND FUTURE RESEARCH

As with any research, limitations present in this study leave room for future research. As with most studies, only a subset of possible conditions were investigated. Only a two-level model (students nested within schools) was investigated. As such, identifiers would only have the potential of being missing at level two. In some cases, researchers would like to investigate models that include additional levels of clustering. The researcher could be interested in students nested in classrooms nested in schools resulting in a three-level model. In this scenario, level two (classroom) and level three (school) identifiers could be missing. Hill and Goldstein (1998) proposed a technique for handling a missing level two identifier in a two level cross-classified model where one level two identifier is known. This technique could be extended to a three-level model. There are scenarios, however, in which both the level two and level three identifiers might be unknown at a particular time point. Future research should investigate methods to handle such missing data.

Only two time points were investigated in this study and the missing identifier was restricted to occur at the first time point as it was assumed that the outcome was measured at the final time point. It is more likely that a longitudinal study would involve more than two time points. As a result, one or more identifiers might be missing for a student. Future research could investigate how procedures for handling missing identifiers perform in scenarios such as this.

In this study, only level one predictors were included and they were modeled as fixed. It is more likely that predictors would be included at both level one and level two and could be a combination of fixed and randomly varying. Future research could investigate the parameter recovery for more complex models including additional predictors as well as predictors at more than one level. Also, the level one predictors used in the present study were all group mean centered around the school's mean at the last time point. There could be a better way to incorporate the set of schools attended (for mobile students) when using group mean centering. As such, this could also provide an additional avenue for research.

As evidenced by the decrease in the relative parameter bias for the *M* coefficient when the true value of the parameter was changed from -0.5 to -5, other coefficient values could be investigated to extend this study's assessment of the impact of parameter values on parameter recovery. Additionally, when using MCMC estimation, one chain was run with 50,000 iterations and a burn-in of 5,000. More stable estimates could result from an increase in iterations and/or burn-in. As such, future research could see if this increase helps stabilize the parameter estimates and also possibly resulting in better recovery of *SE* estimates. Finally, to date, no in-depth research has investigated the minimum sample size needed at each level for acceptable parameter recovery when using MMREM. As such, this is another possible avenue for future research.

IMPORTANCE OF RESEARCH AND CONCLUSION

Real-world datasets frequently contain missing data. In education, a typical higher level clustering unit is school membership. Students change schools for a variety of reasons and, as a result, will move into or out of a study's focal area. As such, it is imperative that researchers understand the most appropriate methods for handling such missingness. This study investigated three ad hoc procedures for handling missing level two units in a two-level model, namely, MMREM-*Unique*, MMREM-*Delete*, and HLM-*Delete*. MMREM-*Unique* assigned a pseudo-level two unit for each missing unit while the other two procedures removed students from the analysis at varying degrees.

MMREM-*Delete* removed students that were missing an identifier while HLM-*Delete* removed all mobile students (students that attended more than one school).

Relative parameter and standard error bias was found across procedures and conditions investigated (percent of mobile students: 10%, 20%; ICC: 15%, 25%; number of level two units: 50, 100; and percent of mobile students missing a level two identifier: 25%, 50%). The procedures that resulted in the least amount of substantial relative parameter bias were MMREM-*Unique* and MMREM-*Delete* while the procedure that resulted in the least amount of substantial relative standard error bias was HLM-*Delete*. Given that parameter estimates' values are typically of most interest to applied researchers, use of MMREM-*Unique* or MMREM-*Delete* is recommended when level two identifiers are missing. Future research should be conducted to continue this line of research to help identify optimal procedures for handling missing identifiers.

References

Adduci, L. L. (1990). *Mobility and student achievement in Orange High School*. Washington, DC: Office of Educational Research and Improvement.

Astone, N. M., & Mclanahan, S. S. (1994). Family structure, residential mobility, and school dropout: A research note. *Demography*, *31(4)*, 575-584.

Barreto, H, & Howland, F. M. (2005). *Introductory econometrics: Using Monte Carlo simulation with Microsoft Excel*. Cambridge University Press.

Been, V., Gould Ellen, I., Schwartz, A. E., Stiefel, L., & Weinstein, M. (2011). Does losing your home mean losing your school?: Effects of foreclosure on the school mobility of children. *Regional Science and Urban Economics*, *41*, 407-414.

Beretvas, S. N. (2010). Cross-classified and multiple membership random effects models. In J. Hox & J. K. Roberts (Eds.), *The handbook of advanced multilevel analysis* (pp. 313-334). New York, NY: Routledge.

Browne, W. J. (2009). *MCMC Estimation in MLwiN, v2.13*. Centre for Multilevel Modeling, University of Bristol.

Browne, W. J., & Draper, D. (2006). A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Biostatistics, 1*, 473-514.

Chandola, T., Clarke, P., Wiggins, R. D., & Bartley, M. (2005). Who you live with and where you live: Setting the context for health using multiple membership multilevel models. *Journal of Epidemiology and Community Health*, *59(2)*, 170-175.

Chung, H., & Beretvas, S. N. (2012). The impact of ignoring multiple membership data structures in multilevel models. *British Journal of Mathematical and Statistical Psychology*, *65(2)*, 185-200.

Clarke, K. A. (2005). The phantom menace: Omitted variable bias in econometrics research. *Conflict Management and Peace Science*, *22*, 341-352.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (revised edition). New York: Academic Press.

Corr, M. C. (1982). A study of the relationship between mobility, academic achievement, and other selected variables in an ethnically diverse school district. (Unpublished doctoral dissertation). Pepperdine University, Malibu, CA.

Cramer, W., & Dorsey, S. (1970). Are movers losers?. *The Elementary School Journal*, *70(7)*, 387-390.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12(2)*, 121-138.

Fielding, A., & Goldstein, H. (2006). *Cross-classified and multiple membership structures in multilevel models: An introduction and review* (Research Report Number 791). Birmingham, UK: University of Birmingham, Department of Education and Skills.

Goldstein, H. (2010). *Multilevel statistical models* (4th ed.). New York: Hodder Arnold.

Goldstein, H., Burgess, S., & McConnell, B. (2007). Modelling the effect of pupil mobility on school differences in educational achievement. *Journal of the Royal Statistical Society, A*., *170(4)*, 941-954.

Greene, W. H. (1993). *Econometric Analysis* (2nd ed.). Macmillan.

Gruman, D. H., Harachi, T. W., Abbott, R. D., Catalano, R. F., & Fleming, C. B. (2008). Longitudinal effects of student mobility on three dimensions of elementary school engagement. *Child Development*, *79(6)*, 1833-1852.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Disruption versus Tiebout improvement: The costs and benefits of switching schools. *Journal of Public Economics*, *88*, 1721-1746.

Hill, P., & Goldstein, H. (1998). Modeling of educational data with cross-classification and missing identification for units. *Journal of Educational and Behavioral Statistics*, *23(2)*, 117-128.

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: an overview and a meta-analysis. *Sociological Methods & Research*, *26(3)*, 329-367.

Kirk, R. E. (1995). *Experimental Design: Procedures for the Behavioral Sciences (3rd ed.)*. New York: Brooks/Cole Publishing.

Kruschke, J. K. (2011). *Doing Bayesian Data Analysis: A Tutorial with R and BUGS*. Burlington, MA: Academic Press.

Leckie, G. (2009). The complexity of school and neighbourhood effects and movements of pupils on school differences in models of educational achievement. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172(3),* 537-554.

Luo, W., & Kwok, O. (2012). The consequences of ignoring individuals' mobility in multilevel growth models: A Monte Carlo study. *Journal of Educational and Behavioral Statistics*, *37(1)*, 31-56.

Mantzicopoulos, P., & Knutson, D. (2000). Head start children: School mobility and achievement in the early grades. *Journal of Educational Research*, *93(5)*, 305-311.

Mehana, M., & Reynolds, A. J. (2004). School mobility and achievement: A meta-analysis. *Children and Youth Services Review*, *26*, 93-119.

Meyers, J. L, & Beretvas, S. N. (2006). The impact of inappropriate modeling of cross-classified data structures. *Multivariate Behavioral Research*, *41(4)*, 473-497.

Rasbash, J. L., & Browne, W. J. (2001). Modeling non-hierarchical structures. In Leyland, A. H. and Goldstein, H. (Eds.), *Multilevel modeling of health statistics* (pp. 93-105). Chichester: John Wiley & Sons.

Rasbash, J., Charlton, C., Browne, W. J., Healy, M., & Cameron, B. (2011). *MLwiN Version 2.24*. Centre for Multilevel Modelling, University of Bristol.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications, Inc.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International, Inc.

Spybrook, J., & Raudenbush, S. W. (2009). An examination of the precision and technical accuracy of the first wave of group-randomized trials funded by the Institute of Education Sciences. *Educational Evaluation and Policy Analysis*, *31*, 298-318.

Strand, S., & Demie, F. (2006). Pupil mobility, attainment and progress in primary school. *British Educational Research Journal*, *32(4)*, 551-568.

Strand, S., & Demie, F. (2007). Pupil mobility, attainment and progress in secondary school. *Educational Studies*, *33(3)*, 313-331.

Temple, J. A., & Reynolds, A. J. (1999). School mobility and achievement: Longitudinal findings from an urban cohort. *Journal of School Psychology*, *37(4)*, 355-377.

Tucker, J., Marx, J., & Long, L. (1998). "Moving on": Residential mobility and children's school lives. *Sociology of Education*, *71(2)*, 111-129.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, *91(3)*, 461-481.

Wolff Smith, L. J., & Beretvas, S. N. (2011, April). *A comparison of techniques for handling and assessing the influence of mobility on student achievement*. Paper presented at the meeting of the American Educational Research Association, New Orleans, Louisiana.

Wolff Smith, L. J., & Beretvas, S. N. (2012, April). *The impact of using incorrect weights with the multiple membership random effects model*. Paper presented at the meeting of the American Educational Research Association, Vancouver, British Columbia, Canada.

Wright, D. (1999). Student mobility: A negligible and confounded influence on student achievement. *The Journal of Educational Research*, *92(6)*, 347-353.